AD-A198 820

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

UNITED STATES AIR FORCE RESEARCH INITIATION PROGRAM

CONDUCTED BY UNIVERSAL ENERGY SYSTEMS UES

Reproduced From
Best Available Copy

1986

TECHNICAL REPORT

VOLUME 1 OF 3

RODNEY C. DARRAH
PROGRAM DIRECTOR, UES

SUSAN K. ESPY
PROGRAM ADMINISTRATOR, UES

LT. COL. CLAUDE CAVENDER
PROGRAM MANAGER, AFOSR

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

DTIC
ELE
AUG 1 6 1988

Best
Available
Copy

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 88-0720
6a. NAME OF PERFORMING ORGANIZATION UNIVERSAL ENERGY SYSTEMS INC.	6b. OFFICE SYMBOL (If applicable)
	7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research/XOT
6c. ADDRESS (City, State, and ZIP Code) 4401 Dayton Xenia Rd Dayton OH 45432	7b. ADDRESS (City, State, and ZIP Code) Building 410 Bolling AFB DC 20332
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR	8b. OFFICE SYMBOL (If applicable) XOT
	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620-85-C-0013
8c. ADDRESS (City, State, and ZIP Code) Building 410 Bolling AFB DC 20332	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 61102F; PROJECT NO. 3396; TASK NO. D5; WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)

USAF Research Initiation Program Volume 1

12. PERSONAL AUTHOR(S)

Program Director Rodney C. Darrah

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Interim	FROM _____ TO _____	April 1988	

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

(SEE REVERSE)

DTIC
ELECTE
AUG 16 1988
S D
D

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Lt. Col. Claude Cavender, Program Manager	22b. TELEPHONE (Include Area Code) (202) 767-4970
	22c. OFFICE SYMBOL XOT

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE

UNITED STATES AIR FORCE

1986 RESEARCH INITIATION PROGRAM

Conducted by

UNIVERSAL ENERGY SYSTEMS, INC.

under

USAF Contract Number F49620-85-C-0013

RESEARCH REPORTS

VOLUME I OF III



Submitted to

Air Force Office of Scientific Research

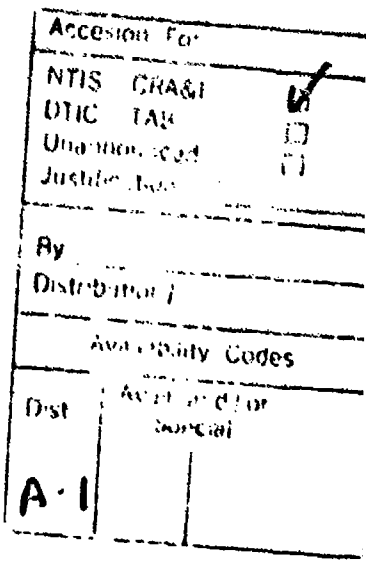
Bolling Air Force Base

Washington, DC

By

Universal Energy Systems, Inc.

April 1988



INTRODUCTION

Research Initiation Program - 1985

AFOSR has provided funding for follow-on research efforts for the participants in the Summer Faculty Research Program. Initially this program was conducted by AFOSR and popularly known as the Mini-Grant Program. Since 1983 the program has been conducted by the Summer Faculty Research Program (SFRP) contractor and is now called the Research Initiation Program (RIP). Funding is provided to establish RIP awards to about half the number of participants in the SFRP.

Participants in the 1985 SFRP competed for funding under the 1985 RIP. Participants submitted cost and technical proposals to the contractor by 1 November 1985, following their participation in the 1985 SFRP.

Evaluation of these proposals was made by the contractor. Evaluation criteria consisted of:

1. Technical Excellence of the proposal
2. Continuation of the SFRP effort
3. Cost sharing by the University

The list of proposals selected for award was forwarded to AFOSR for approval of funding. Those approved by AFOSR were funded for research efforts to be completed by 31 December 1986.

The following summarizes the events for the evaluation of proposals and award of funding under the RIP.

A. Rip proposals were submitted to the contractor by 1 November 1985. The proposals were limited to $20,000 plus cost sharing by the universities. The universities were encouraged to cost share since this is an effort to establish a long term effort between the Air Force and the university.

B. Proposals were evaluated on the criteria listed above and the final award approval was given by AFOSR after consultation with the Air Force Laboratories.

C. Subcontracts were negotiated with the universities. The period of performance of the subcontract was between October 1985 and December 1986.

Copies of the Final Reports are presented in Volumes I through III of the 1985 Research Initiation Program Report. There were a total of 82 RIP awards made under the 1985 program.

Reproduced From
Best Available Copy

TABLE OF CONTENTS

SECTION	PAGE
INTRODUCTION	i
STATISTICS	ii
PARTICIPANT LABORATORY ASSIGNMENT	vii
RESEARCH REPORTS	xv

INTRODUCTION

Research Initiation Program - 1986

AFOSR has provided funding for follow-on research efforts for the participants in the Summer Faculty Research Program. Initially this program was conducted by AFOSR and popularly known as the Mini-Grant Program. Since 1983 the program has been conducted by the Summer Faculty Research Program (SFRP) contractor and is now called the Research Initiation Program (RIP). Funding is provided to establish RIP awards to about half the number of participants in the SFRP.

Participants in the 1986 SFRP competed for funding under the 1986 RIP. Participants submitted cost and technical proposals to the contractor by 1 November 1986, following their participation in the 1986 SFRP.

Evaluation of these proposals was made by the contractor. Evaluation criteria consisted of:

1. Technical Excellence of the proposal
2. Continuation of the SFRP effort
3. Cost sharing by the University

The list of proposals selected for award was forwarded to AFOSR for approval of funding. Those approved by AFOSR were funded for research efforts to be completed by 31 December 1987.

The following summarizes the events for the evaluation of proposals and award of funding under the RIP.

A. Rip proposals were submitted to the contractor by 1 November 1986. The proposals were limited to $20,000 plus cost sharing by the universities. The universities were encouraged to cost share since this is an effort to establish a long term effort between the Air Force and the university.

B. Proposals were evaluated on the criteria listed above and the final award approval was given by AFOSR after consultation with the Air Force Laboratories.

C. Subcontracts were negotiated with the universities. The period of performance of the subcontract was between October 1986 and December 1987.

Copies of the Final Reports are presented in Volumes I through III of the 1986 Research Initiation Program Report. There were a total of 98 RIP awards made under the 1986 program.

STATISTICS

Total SFRP Participants	158
Total RIP Proposals submitted by SFRP	134
Total RIP Proposals submitted by GSSSP	7
Total RIP Proposals submitted	141

Total RIP's funded to SFRP	94
Total RIP's funded to GSSSP	4
Total RIP's funded	98

Total RIP's Proposals submitted by HBCU's	14
Total RIP's Proposals funded to HBCU's	9

Laboratory	SFRP Participants	RIP's Submitted	RIP's Funded
AAMRL	9	11 (3 GSSSP)	5 (1 GSSSP)
APL	8	7	6
-AD	11	12 (3 GSSSP)	9 (3 GSSSP)
AEDC	6	6	3
-AL	7	6	5
BRMC	2	2	0
LC	1	1	1
ESMC	0	0	0
ESD	2	1	1
ESC	7	6	5
FOL	13	13 (1 GSSSP)	9
FJSRL	8	8	4
GL	13	9	7
HRL/OT	4	4	2
HRL/LR	4	4	2
HRL/MO	3	2	2
HRL/ID	4	4	3
LMC	2	1	1
-ML	11	8	6
OEHL	4	4	3
RPL	5	4	4
RADC	9	8	7
SAM	17	13	8
WHMC	1	1	1
WL	7	6	4
Total	158	141	98

LIST OF UNIVERSITY THAT PARTICIPATED

Adelphi University	1	Meharry Medical College	2
Alabama A&M University	2	Miami University of Ohio	1
Alabama, University of	5	Miami, University of	1
Alaska, University of	1	Mississippi State University	1
Alfred University	2	Mississippi, University of	1
Auburn University	1	Missouri, University of	2
Boise State University	1	Morehouse College	2
Bradley University	1	Motlow State College	1
Brown University	1	Nebraska, University of	1
Carleton College	1	New Mexico, University of	1
Catholic University of America	1	New Orleans, University of	1
Cedarville College	1	New York University	1
Cincinnati, University of	3	Norfolk State University	1
Colorado, University of	1	North Carolina A&T University	1
Dartmouth College	1	North Carolina, University of	1
Davidson College	1	North Texas State University	1
Dayton, University of	5	Northern Arizona Univeristy	1
Drexel University	1	Northwestern University	1
Duke University	1	Oakwood College	1
Eastern Kentucky University	1	Ohio State University	4
Eastern Montana College	1	Ohio University	3
Edinboro University	1	Oklahoma State University	2
Florida Atlantic University	1	Oklahoma, University of	1
Florida International University	1	Oregon State University	1
Florida State University	2	Pacific University	1
Florida University	3	Paine College	1
Florida, University of	3	Pennsylvania State University	1
Franklin and Marshall College	1	Portland, University of	1
Georgia Institute of Technology	1	Purdue University	2
Georgia, University of	2	Scanton, University of	1
Grambling State University	1	South Carolina, University of	1
Houghton College	1	Southern Illinois University	1
Indiana University	2	Southern Michigan, University of	1
Iowa State University	1	Southern University	1
Iowa, University of	1	Stetson University	1
Jackson State University	4	Stevens Institute of Technology	1
Jefferson State	1	Syracuse, University of	1
Jesm Baromedical Research	1	Tennessee, University of	1
Kansas State Unviersity	1	Texas A&I University	1
Kennesaw University	1	Texas A&M University	2
Lehigh University	1	Texas Southern University	2
Louisiana State University	5	Texas, University of	2
Lowell, University of	2	The Citadel	2
Lyndon State College	1	Toledo, University of	1

(Continued)

LIST OF UNIVERSITY THAT PARTICIPATED

MIT	1	Touglaoo College	1
Maine, University of	1	Trinity University	1
Marquette University	1	Tulsa, University of	2
Mary Washington College	1	U.S. Naval Academy	1
Massachusettes, University of	1	Valparaiso University	1
Vanderbilt University	1	West Virginia University	1
Warren Wilson College	1	Wichita State University	1
Washington State University	2	Wisconsin-Eau Claire	1
Wayne State University	1	Worchester Polytechnic	1
West Florida, University of	1	Wright State University	4
West Georgia College	1	Wyoming, University of	2
		Xavier University	1

PARTICIPANTS LABORATORY ASSIGNMENT

PARTICIPANT LABORATORY ASSIGNMENT (Page 1)

ARMAMENT LABORATORY
(Eglin Air Force Base)

Dr. Prabhat Hajela
University of Flordia
Specialty: Aeronautics & Astronautics

Dr. Boghos D. Sivazlian
The University of Florida
Specialty: Operations Research

Dr. David I. Lawson
Stetson University
Specialty: Mathematics

Mr. Chris Reed (GSRP)
Florida University
Specialty: Aerodynamics

Dr. Barbara Rice
Alabama A&M University
Specialty: Mathematics

Mr. Jim Sirkis (GSRP)
Florida University
Specialty: Engineering Mechanics

Dr. Sally A. Sage
West Georgia College
Specialty: Computer Science

Ms. Jennifer Davidson (GSRP)
Florida University
Specialty: Mathematics

Dr. Meckinley Scott
University of Alabama
Specialty: Statistics

ARNOLD ENGINEERING DEVELOPMENT CENTER
(Arnold Air Force Systems)

Dr. Glen Johnson
Vanderbilt University
Specialty: Mechanical Eng.

Dr. Arthur A. Mason
The University of Tennessee
Specialty: Physics

ELECTRONIC SYSTEMS DIVISION
(Hanscom Air Force Base)

Dr. Stephan E. Kolitz
University of Massachusetts
Specialty: Operations Research

PARTICIPANT LABORATORY ASSIGNMENT (Page 2)

ENGINEERING AND SERVICES CENTER
(Tyndall Air Force Base)

Dr. Thomas A. Carney
Florida State University
Specialty: Meteorology

Dr. William T. Cooper
Florida State University
Specialty: Chemistry

Dr. Yong S. Kim
The Catholic Univ. of America
Specialty: Civil Engineering

Dr. Cheng Liu
University of North Carolina
Specialty: Civil Engineering

Dr. Roy M. Ventullo
University of Dayton
Specialty: Microbiology

FRANK J. SEILER RESEARCH RESEARCH LABORATORY
(United State Air Force Academy)

Dr. David R. Anderson
University of Colorado
Specialty: Organic Chemistry

Dr. Bernard J. Piersma
Houghton College
Specialty: Physical Chemistry

Dr. William D. Siuru, Jr.
University of Colorado
Specialty: Mechanical Eng.

Dr. Timothy R. Troutt
Washington State University
Specialty: Mechanical Eng.

GEOPHYSICS LABORATORY
(Hanscom Air Force Base)

Dr. John E. Ahlquist
Florida State University
Specialty: Meteorology

Dr. Frank P. Battles
Mass. Maritime Academy
Specialty: Physics

Dr. Wolfgang Christian
Davidson College
Specialty: Physics

Dr. Donald F. Collins
Warren Wilson College
Specialty: Physics

Dr. Patrick T. Gannon, Sr.
Lyndon State College
Specialty: Atmospheric Science

Dr. C. Randal Lishawa
Jefferson State University
Specialty: Physical Chemistry

Dr. Robert M. Nehs
Texas Southern University
Specialty: Mathematics

PARTICIPANT LABORATORY ASSIGNMENT (Page 3)

LOGISTICS COMMAND
(Wright-Patterson Air Force Base)

Dr. Ming-Shing Hung
Kent State University
Specialty: Business Administration
Management Science

LOGISTICS MANAGEMENT CENTER
(Gunter Air Force System)

Dr. Dan B. Rinks
Louisiana State University
Specialty: Quantitative Mgmt. Science

ASTRONAUTICS LABORATORY
(Edwards Air Force Base)

Dr. William M. Grissom
Morehouse College
Specialty: Mechanical Engineering

Dr. Joel R. Klink
Univ. of Wisconsin-Eau Claire
Specialty: Organic Chemistry

Dr. Siavash H. Sohrab
Northwestern University
Specialty: Engineering Physics

Dr. Nicholas E. Takach
University of Tulsa
Specialty: Chemistry

ROME AIR DEVELOPMENT CENTER
(Griffis Air Force Base)

Dr. Donald F. Hanson
University of Mississippi
Specialty: Electrical Engineering

Dr. John M. Jobe
Miami University of Ohio
Specialty: Statistics

Dr. Philipp G. Kornreich
Syracuse University
Specialty: Electrical Engineering

Dr. Mou-Liang Kung
Norfolk State University
Specialty: Mathematics

Dr. Craig G. Prohazka
University of Lowell
Specialty: Electrical Engineering

Dr. Richard S. Quimby
Worcester Polytechnic Institute
Specialty: Physics

Dr. Stephen T. Welstead
University of Alabama
Specialty: Applied Mathematics

PARTICIPANT LABORATORY ASSIGNMENT (Page 4)

WEAPONS LABORATORY
(Kirtland Air Force Base)

Dr. Albert W. Biggs
University of Alabama
Specialty: Electrical Eng.

Dr. Alexandru A. Pelin
Florida International Univ.
Specialty: Computer Science

Dr. Fabian C. Hadipriono
The Ohio State University
Specialty: Engineering, Civil

Dr. Martin A. Shadday, Jr.
University of South Carolina
Specialty: Mechanical Engineering

AERO PROPULSION LABORATORY
(Wright-Patterson Air Force Base)

Dr. Lea D. Chen
The University of Iowa
Specialty: Organic Chemistry

Dr. James C. Ho
Wichita State University
Specialty: Chemistry

Dr. Jacob N. Chung
Washington State University
Specialty: Mechanical Engineering

Dr. Mo Samimy
Ohio State University
Specialty: Mechanical Engineering

Dr. Shirshak K. Dhali
Southern Illinois University
Specialty: Electrical Engineering

Dr. Robert P. Taylor
Mississippi State University
Specialty: Mechanical Engineering

AVIONICS LABORATORY
(Wright-Patterso Air Force Base)

Dr. John Y. Cheung
University of Oklahoma
Specialty: Electrical Engineering

Dr. Dennis W. Whitson
Indiana Univ. of Pennsylvania
Specialty: Physics

Dr. William A. Grosky
Wayne State University
Specialty: Eng. & Applied Science

Dr. George W. Zobrist
University of Missouri-Rolla
Specialty: Electrical Engineering

Dr. Ken Tomiyama
Pennsylvania State University
Specialty: System Science

PARTICIPANT LABORATORY ASSIGNMENT (Page 5)

FLIGHT DYNAMICS LABORATORY
(Wright-Patterson Air Force Base)

Dr. Bor-Chin Chang
Bradley University
Specialty: Electrical Eng.

Dr. George R. Doyle, Jr.
University of Dayton
Specialty: Mechanical Eng.

Dr. Paul S.T. Lee
N.C. A&T State University
Specialty: Quantitative Methods

Dr. V. Dakshina Murty
University of Portland
Specialty: Engineering Mechanics

Dr. Singiresu S. Rao
Purdue University
Specialty: Engineering Design

Dr. Tsun-wai G. Yip
Ohio State University
Specialty: Aeronautics

Dr. Ajmal Yousuff
Drexel University
Specialty: Aeronautics

Dr. Richard W. Young
University or Cincinnati
Specialty: Applied Mechanics

Dr. Peter J. Disimile
University of Cincinnati
Specialty: Fluid Mechanics

MATERIALS LABORATORY
(Wright-Patterson Air Force Base)

Dr. Lokesh R. Dharani
University of Missouri-Rolla
Specialty: Engineering Mechanics

Dr. Gerald R. Graves
Louisiana State University
Specialty: Industrial Engineering

Dr. Gopal M. Mehrotra
Wright State University
Specialty: Metallurgy

Dr. Robert A. Patsiga
Indiana Univ. of Pennsylvania
Specialty: Organic Polymer Chem.

Dr. Nisar Shaikh
University of Nebraska-Lincoln
Specialty: Applied Mathematics

Dr. Stuart R. Stock
Georgia Institute of Technology
Specialty: Metallurgy

PARTICIPANT LABORATORY ASSIGNMENT (Page 6)

HARRY G. ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY
(Wright-Patterson Air Force Base)

Ms. Beverly Girten
Ohio University
Specialty: Physiology

Dr. Albert R. Wellens
University of Miami
Specialty: Experimental Social

Dr. Jacqueline G. Paver
Duke University
Specialty: Biomechanical Eng.

Dr. Robert L. Yolton
Pacific University
Specialty: Psychology, Optometry

Dr. Kuldip S. Rattan
Wright State University
Specialty: Electrical Engineering

HUMAN RESOURCES LABORATORY - LOGISTICS AND HUMAN FACTORS DIVISION
(Wright-Patterson Air Force Base)

Dr. Patricia T. Boggs
Wright State University
Specialty:Decision Science

Dr. Stephen L. Loy
Iowa State University
Specialty: Management Information

HUMAN RESOURCES LABORATORY - OPERATIONS TRAINING DIVISION
(Williams Air Force Base)

Dr. Billy R. Wooten
Brown University
Specialty:Philosophy, Psychology

HUMAN RESOURCES LABORATORY - MANPOWER AND PERSONNEL DIVISION
(Brooks Air Force Base)

Dr. Richard H. Cox
Kansas State University
Specialty:Motor Learning & Control

Dr. Jorge L. Mendoza
Texas A&M University
Specialty: Psychology

USAF OCCUPATIONAL AND ENVIRONMENT HEALTH LABORATORY
(Brooks Air Force Base)

Dr. Clifford C. Houk
Ohio University
Specialty: Inorganic Chemistry

Dr. Shirley A. Williams
Jackson State University
Specialty: Physiology

Dr. Ralph J. Rascati
Kennesaw College
Specialty: Biochemistry

HUMAN RESOURCES LABORATORY - TRAINING SYSTEMS
(Brooks Air Force Base)

Dr. Charles E. Lance
University of Georgia
Specialty: Psychology

Dr. Doris J. Walker-Dalhouse
Jackson State University
Specialty: Reading Education

Dr. Philip D. Olivier
University of Texas
Specialty: Electrical Engineering

SCHOOL OF AEROSPACE MEDICINE
(Brooks Air Force Base)

Dr. Hoffman H. Chen
Grambling State University
Specialty: Mechanical Engineering

Dr. Frank O. Hadlock
Florida Atlantic University
Specialty: Mathematics

Dr. Brenda J. Claiborne
University of Texas
Specialty: Biology

Dr. Parsottam J. Patel
Meharry Medical College
Specialty: Microbiology

Dr. Vito G. DelVecchio
University of Scranton
Specialty: Biochemistry, Genetics

Dr. Richard M. Schori
Oregon State University
Specialty: Mathematics

Dr. Ramesh C. Gupta
University of Maine at Orono
Specialty: Mathematical Statistics

Dr. Shih-sung Wen
Jackson State University
Specialty: Educational Psychology

WILFORD HALL MEDICAL CENTER
(Lackland Air Force Base)

Dr. Donald W. Welch
Texas A&M University
Specialty: Microbiology

RESEARCH REPORTS

MINI-GRANT RESEARCH REPORTS of the A.F.

1986 RESEARCH INITIATION PROGRAM in Vol. 1

include:

Technical Report Number	Title and Mini-Grant No.	Professor
Volume I		
1	Weather Forecast Evaluation be Decomposition of the Wind Field into Barotropic and Baroclinic Components; 760-6MG-041	Dr. Jon Ahlquist
2	An EPR Study of the Role of Catalysts in the Thermal Decomposition of Nitroaromatic Compounds; 760-6MG-044	Dr. David Anderson
3	Stellar Scintillometer Based Studies of Optical Turbulence; 760-6MG-058	Dr. Frank Battles
4	Requested A No-Cost Time Extention. To Be Submitted In 1987 Mini-Grant Final Report. 760-6MG-072	Dr. Albert Biggs
5	Basic Research on the Impact of Cognitive Styles on Decision Making; 760-6MG-127	Dr. Patricia Boggs
6	A Feasibility Study and Test Application of Uncertainty Estimates to an Atmospheric Dispersion Model with Potential Utility in Air Force Operations 760-6MG-050	Dr. Thomas Carney
7	Design of H Multivariable Optimal Control Systems; 760-6MG-013	Dr. Bor-Chin Chang
8	Visualization of Hydrocarbon Jet Diffusion Flames; 760-6MG-113	Dr. Lea Chen

9	Requested A No-Cost Time Extention. To Be Submitted In 1987 Mini-Grant Final Report. 760-6MG-118	Dr. Hoffman Chen
10	Report Not Received In Time. Will Be Provided When Available. 760-6MG-135	Dr. John Cheung
11	Infrared Fluorescense and Photo-acoustic Measurements of NO (v=2) Relaxation as a Function of Temperature ; 760-6MG-030	Dr. Wolfgang Christian
12	Heat and Mass Transfer in a Dual-Latent Heat Packed Bed Thermal Storage System ; 760-6MG-067	Dr. Jacob Chung
13	Long-term Potentiation in Interneurons in the Mammalian Brain. 760-6MG-101	Dr. Brenda Claiborne
14	The Development of Image Processing Algorithms for AFGL Ultraviolet Camera and Other Imaging Systems. 760-6MG-028	Dr. Donald Collins
15	Report Not Received In Time. Will Be Provided When Available. 760-6MG-081	Dr. William Cooper
16	Relationship Between Stages of Motor Learning and Kinesthetic Sensitivity ; 760-6MG-069	Dr. Richard Cox
17	Received A No-Cost Time Extention. To Be Submitted In 1987 Mini-Grant Final Report. 760-6MG-024	Ms. Jennifer Davidson
18	Report Not Received In Time. Will Be Provided When Available. 760-6MG-076	Dr. Vito DelVecchio

13 cont

19	Investigation of Pulsed Discharges in Nitrogen for Plasma Processing; 760-6MG-046	Dr. Shirshak Dhali
20	Modeling of Failure Mechanisms in Ceramic Composites Under Flexural Loading; 760-6MG-115	Dr. Lokesh Dharani
21	Requested A No-Cost Time Extention. To Be Submitted In 1987 Mini-Grant Final Report. 760-6MG-075	Dr. Peter Disimile
22	Requested A No-Cost Time Extention. To Be Submitted In 1987 Mini-Grant Final Report. 760-6MG-006	Dr. George Doyle
23	Sensitivity of Mesoscale Wind to Variations in Vegetation Canopy Parameters and Surface Properties; 760-6MG-100	Dr. Patrick Gannon
24	Effects of Exercise and Dobutamine on Suspension Hypokinesia/ Hypodunamia Deconditioning in Rats; 760-6MG-139	Ms. Beverly Girten
25	An Investigation of Computer Communications Using Knowledge-Based Systems; 760-6MG-015	Dr. Gerald Graves
Volume II		
26	Droplet Size Distribution Measurement In A Single Element Liquid Rocket Injector. (edc) 760-6MG-040	Dr. William Grissom
27	A Unified Approach of the Linear Camera Calibration Problem 760-6MG-070	Dr. William Grosky
28	Survival Anaysis of Radiated Animals for Small Sample Sizes 760-6MG-053	Dr. Ramesh Gupta
29	Report Not Received In Time. Will Be Provided When Available. 760-6MG-054	Dr. Fabian Hadipriono

30	Requested A No-Cost Time Extention. To Be Submitted In 1987 Mini-Grant Final Report. 760-6MG-073	Dr. Frank Hadlock
31	Studies in Optimum Shape Synthesis for Structures Undergoing Plastic Deformation 760-6MG-002	Dr. Prabhat Hajela
32	Report Not Received In Time. Will Be Provided When Available. 760-6MG-092	Dr. Donald Hanson
33	Pulsed Power Conductors 760-6MG-005	Dr. James Ho
34	The Locally Implicit Method for Computational Aerodynamics 760-6MG-111	Dr. Peter Hoffman
35	Fluorescent Dye Binding Identification of Asbestos on Membrane Filters and in Bulk Materials 760-6MG-066	Dr. Clifford Houk
36	Requested A No-Cost Time Extention. To Be Submitted In 1987 Mini-Grant Fianl Report. 760-6MG-105	Dr. Ming S. Hung
37	Report Not Received In Time. Will Be Provided When Available. 760-6MG-019	Dr. John Jobe
38	Expert System for Optimal Design 760-6MG-016	Dr. Glen Johnson
39	Report Not Received In Time. Will Be Provided When Available. 760-6MG-004	Dr. Yong Kim
40	The Synthesis of Some New Energetic Materials 760-6MG-056	Dr. Joel Klink

41	Report Not Received In Time. Will Be Provided When Available. 760-6MG-094	Dr. Steve Kolitz
42	MBE Grown Al-Cu Alloy Films 760-6MG-090	Dr. Philipp Kornreich
43	Simulation for Priority Handling Algorithms 760-6MG-011	Dr. Mou-Liang Kung
44	Received A No-Cost Time Extention. To Be Submitted In 1987 Mini-Grant Final Report. 760-6MG-031	Dr. Charles Lance
45	A Neural Network Simulation Generator, Simualtions of Learned Serial Behavior, and a Neural Explanation of Emergent Communication 760-6MG-001	Dr. David Lawson
46	Data Processing and Statistical Analysis of In-Service Aircraft Transparency Failures 760-6MG-023	Dr. Paul Lee
47	Trajectory Studies of the Bimolecular Reaction of H2Ov/H2O 760-6MG-107	Dr. C. Lishawa
48	Comparison of Field Rut Depth Measurements and Rutting Predicitions of Asphalt Pavement Under High Tire Pressure and Temperature 760-6MG-009 *NOT PUBLISHABLE AT THIS TIME*	Dr. Cheng Liu
49	Report Not Received In Time. Will Be Provided When Available. 760-6MG-134	Dr. Stephen Loy
50	Received A No-Cost Time Extention. To Be Submitted In 1987 Mini-Grant Final Report. 760-6MG-099	Dr. Arthur Mason

51	Report Not Received In Time. Will Be Provided When Available. 760-6MG-121	Dr. Gopal Mehrotra
52	Report Not Received In Time. Will Be Provided When Available. 760-6MG-136	Dr. Jorge Mendoza
53	Report Not Received In Time. Will Be Provided When Available. 760-6MG-079	Dr. Dakshina Murty
54	Development of a New Finite Element Grid for Limites Area Weather Models 760-6MG-120	Dr. Robert Nehs
55	Report Not Received In Time. Will Be Provided When Available. 760-6MG-032	Dr. Philip Olivier
56	Report Not Received In Time. Will Be Provided When Available. 760-6MG-131	Dr. Parsottam Patel
57	Report Not Received In Time. Will Be Provided When Available. 760-6MG-065	Dr. Robert Patsiga
58	Report Not Received In Time. Will Be Provided When Available. 760-6MG-020	Dr. Jacqueline Paver
59	Automatic Program Generation from Specifications Using Prolong 760-6MG-117	Dr. Alexandru Pelin
60	Some Novel Aspects of Organic Electrochemistry in Room Temperature Molten Salts 760-6MG-038	Dr. Bernard Piersma
61	Improved Distributed Operating System Communication Protocols 760-6MG-061	Dr. Craig Prohazka

62	Tunable Infrared to Visible Light Conversion in Rare Earth and Transition Metal Doped Fluoride Glassess 760-6MG-042	Dr. Richard Quimby
63	Optimal Structural Modifications to Enhance the Robustness of Actively Controlled Large Flexible Structures 760-6MG-036	Dr. Singiresu Rao
64	Report Not Received In Time. Will Be Provided When Available. 760-6MG-062	Dr. Ralph Rascati
65	MATRIXs-Based Computer Simulation of the Cardiovascular System Under +Gz Stress 760-6MG-104	Dr. Kuldip Rattan
66	Adaptive Grid Generation Techniques for TRansonic Projectile Base Flow Problems 760-6MG-034	Mr. Chris Reed

Volume III

67	Utilization of the Image Algebra 760-6MG-106	Dr. Barbara Rice
68	Simulation Studies of MICAP Allocation Systems for EOQ Items 760-6MG-084	Dr. Dan Rinks
69	Computer Modeling of Infrared Signatures 760-6MG-017	Dr. Sally Sage
70	Received A No-Cost Time Extention. To Be Submitted In 1987 Mini-Grant Final Report. 760-6MG-059	Dr. Mo Samimy
71	An Intentional Tutor 760-6MG-052	Dr. Richard Schori
72	Report Not Received In Time. Will Be Provided When Available. 760-6MG-025	Dr. Meckinley Scott

73	Report Not Received In Time. Will Be Provided When Available. 760-6MG-089	Dr. Martin Shadday
74	Report Not Received In Time. Will Be Provided When Available. 760-6MG-007	Dr. Nisar Shaikh
75	Report Not Received In Time. Will Be Provided When Available. 760-6MG-142	Mr. Jim Sirkis
76	Two-Dimensional Flight Simulation Model for an Aircraft with a Rapodly Rotating Airfoil 760-6MG-071	Dr. William Siuru
77	Mission Effectiveness Analysis of an Aircraft Attacking Passive Targets 760-6MG-018	Dr. Boghos Sivazlian
78	Requested A No-Cost Time Extention. To Be Submitted In 1987 Mini-Grant Final Report. 760-6MG-110	Dr. Siavash Sohrab
79	Synchrotron White Beam Topography of Striations and Interface Breakdown in GaAs and of Strain Fields in Si 760-6MG-103	Dr. Stuart Stock
80	Received A No-Cost Time Extention. To Be Submitted In 1987 Mini-Grant Final Report. 760-6MG-130	Dr. Nicholas Takach
81	Complete Statistical Classification of Natural Surface Roughness on Gas Turbine Blades 760-6MG-064	Dr. Robert Taylor
82	Evaluation of Atmospheric Effects for Operational Tactical Decision Aid 760-6MG-047	Dr. Ken Tomiyama

83	An Ivestigation Concerning the Formation of a Dynamic Stall Vortex on a Pitching Airfoil 760-6MG-087	Dr. Timothy Troutt
84	Biodegradation of Aqueous Film Forming Foam Components in Laboratory Scale Microcosms 760-6MG-124	Dr. Roy Ventullo
85	Requested A No-Cost Time Extention. To Be Submitted In 1987 Mini-Grant Fianl Report. 760-6MG-080	Dr. Doris Walker-Dalhouse
86	Received A No-Cost Time Extention. To Be Submitted In 1987 Mini-Grant Fianl Report. 760-6MG-091	Dr. Donald Welch
87	Effects of Telecommunication Media upon Group Decision Making Processess within a Multi-Team Situation Assessment Task 760-6MG-085	Dr. Albert Wellens
88	Report Not Received In Time. Will Be Provided When Available. 760-6MG-063	Dr. Steve Welstead
89	Can a supervisory Control Simulation System Assess Cognitive Abilities? 760-6MG-049	Dr. Shih-sung Wen
90	Effects on the BICFET of the Fermi Distribution Factor and the Al Mole Fraction 760-6MG-088	Dr. Dennis Whitson
91	The Warehouse Layout Program 760-0MG-038	Dr. Jesse Williams
92	Received A No-Cost Time Extention. To Be Submitted In 1987 Mini-Grant Final Report. 760-6MG-078	Dr. Shirley Williams

93	Report Not Received In Time. Will Be Provided When Available. 760-6MG-051	Dr. Billy Wooten
94	Report Not Received In Time. Will Be Provided When Available. 760-6MG-109	Dr. Tsun-wai Yip
95	Changes in Perceived Workload and Physiological Responses Associated with Monocular Versus Binocular Viewing Conditions 760-6MG-037	Dr. Robert Yolton
96	Finite Element Analysis of Thermo-mechanically Coupled Stress and Temperature Fields 760-6MG-129	Dr. Richard Young
97	Simplification of H$_\infty$ Compensators 760-6MG-098	Dr. Ajmal Yousuff
98	Late Appointment. Final Report Will Be Provided When Available. 760-6MG-055	Dr. George Zobrist

1655s

1986 USAF-UES RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by

Universal Energy Systems, Inc.

FINAL REPORT

Weather Forecast Evaluation by Decomposition of the Wind Field into Barotropic and Baroclinic Components

Prepared by: Dr. Jon E. Ahlquist

Academic Rank: Associate Professor

Department: Meteorology

University: Florida State University
Tallahassee, FL 32306-3034

USAF Scientific Contact: Dr. Chien-hsiung Yang
Atmospheric Prediction Branch (LYP)
Air Force Geophysics Laboratory
Hanscom Air Force Base

Date: 14 December 1987

Weather Forecast Evaluation by Decomposition of the Wind Field into Barotropic and Baroclinic Components

by

Dr. Jon E. Ahlquist

ABSTRACT

This report is a follow-up to the work performed at the Air Force Geophysics Laboratory during the summer of 1986 which examined the divergent (baroclinic) component of two weather forecasts computed by the Air Force's global spectral model. Here, we focus on the nondivergent (barotropic) component of the planetary circulation for the same two four-day forecasts studied during the summer of 1986.

The accuracy of two four-day forecasts, one for February 1979 and one for June 1979, was examined by projecting forecasts and observations onto planetary-scale barotropic normal modes. The forecast errors are modest and are generally smaller than the errors produced by a statistical forecasting method. Although there is still room for improvement in the AFGL forecasts, the errors in the wind and mass fields at planetary scales appear to be unbiased.

1. Introduction

Daley et al. (1981) showed that one way in which a numerical weather forecast can suffer large errors is through oversimplified tropical initial conditions and/or a model domain which is hemispheric instead of global. When such inadequacies existed, they found that barotropic normal modes were excited which were an order of magnitude stronger than they should have been.

In the present study, we have repeated the basic procedure of Daley et al. (1981). Since the model domain for the Air Force weather prediction model is global and since reasonable tropical initial conditions were used, we did not expect (nor did we find) the gross errors encountered by Daley et al. (1981), but we did want to see what errors still remained in the barotropic component of the forecast and whether there were any discernable biases in the errors at planetary scales. We did find errors in the forecasts, but the errors seem to unbiased.

Section 2 describes the forecast periods, the method for examining forecast error, and a statistical forecasting procedure used as a comparison against the AFGL spectral model. Results are discussed in section 3, and conclusions in section 4. The appendices contain ancillary material.

2. Description of the forecast and analysis procedure

The two forecast records studied herein were computed using the AFGL global spectral model described by Brenner et al. (1984), which used a rhomboidal 30 truncation. The first forecast was based on the NMC global weather anlysis for 1200 UTC, 17 February 1987, and the forecast extended for four days. We examined four once-daily forecasts for 1200 UTC for 18 through 21 February 1987 and compared them against NMC global analyses for the same times. The second forecast was based on the NMC global weather analysis for 0000 UTC, 10 June 1979, and also lasted for four days. We have verified the June forecasts using NMC analyses for 0000 UTC each day.

Before the analyses and forecasts could be decomposed as a linear combination of normal modes, we had to construct the normal modes. We confined our interest to external modes, whose horizontal structures are the same as the normal modes of the shallow water equations. Following Ahlquist (1982), normal modes for the shallow water equations were computed by computer program HNORMODG using a Galerkin method assuming a resting basic state. (See appendix C for a list of software mailed to AFGL.) The normal modes of interest are Rossby waves; gravity wave modes were ignored. The process of projecting observations onto normal modes is described in Ahlquist (1982), with the exception that, in this case, the projection was done individually for each pressure level so that vertical structure would be evident. Projection of the forecasts and analyses onto the normal modes was performed using the computer program HPROJCT2. The results

were time series A(p,t) and B(p,t) for each normal mode, where the component of atmospheric circulation due to each mode is described by

$$u(\varphi,\lambda,p,t) = (A(p,t) \cos m\lambda + B(p,t) \sin m\lambda) (gh_e)^{1/2} U(\varphi)$$

$$v(\varphi,\lambda,p,t) = (-B(p,t) \cos m\lambda + A(p,t) \sin m\lambda) (gh_e)^{1/2} V(\varphi)$$

$$h(\varphi,\lambda,p,t) = (A(p,t) \cos m\lambda + B(p,t) \sin m\lambda) \quad h_e \quad H(\varphi),$$

where:

(φ,λ,p,t) = (latitude, longitude, pressure, time),

m = zonal wavenumber,

g = 9.81 Newtons/kg,

h_e = 10 km = equivalent depth, and

U, V, H = nondimensional functions of latitude called Hough functions.

A, B, U, V, and H are different for each normal mode.

Although forecasts are often evaluated by comparing how well they do against "persistence," we wanted a more realistic standard for forecast comparison. It is well known that barotropic normal modes in the Earth's atmosphere can propagate for days (Lindzen et al., 1984; Ahlquist, 1985), so instead of using "in situ" persistence, we chose to compare model skill against a statistical forecast that allows for propagation and decay. Specifically, we modeled each time series as

$$W(t) = a(1)\, W(t - 1/2 \text{ day}) + a(2)\, W(t - 1 \text{ day}) + E(t)$$

where

$W(t)$ = complex-valued time series representing the departure from the time mean of the wave amplitude and phase
= $[A(t) - \text{avg}(A)] + i\,[B(t) - \text{avg}(B)]$.

$a(1)$, $a(2)$ = complex-valued coefficients

$E(t)$ = complex-valued error.

The coefficients a(1) and a(2) can be computed by a number of methods, basically all of which minimize the magnitude of the average squared error, $|E|^2$. The model in the equation above is called a second order autoregressive (AR) model. Autoregressive modeling is covered extensively for real-valued processes by Box and Jenkins (1976). In appendix 2, we have reproduced from Box and Jenkins the classical derivation of the Yule-Walker method for computing AR coefficients, here modified for complex-valued processes. Subroutines for estimating AR coefficients are contained in the computer files MEMRL and MEMCO for real and complex processes, respectively. (MEM stands for maximum entropy method.) For the present research project, we only used MEMCO, but we have also supplied MEMRL to the AFGL researchers for completeness. The program that called the autoregressive fitting subroutine and performed the statistical forecast is named HFCAST.

In order to use the AR model for statistical forecasting, we proceeded as follows. Let t_o represent the most recent time for which we have observations, and let $w(t_o + j\,dt)$ denote our forecast for W(t) at time $t_o + j\,dt$, where dt is the time increment between observations (and forecasts). Then

$$w(t_o + dt) = a(1)\, W(t_o) + a(2)\, W(t_o - dt)$$

$$w(t_o + 2\,dt) = a(1)\, w(t_o + dt) + a(2)\, W(t_o)$$

$$w(t_o + 3\,dt) = a(1)\, w(t_o + 2\,dt) + a(2)\, w(t_o + dt),$$

etc.

3. Discussion of results

3.1 Preliminaries

Normal mode Rossby waves can be indexed by their zonal wavenumber and by a meridional wavenumber based on the number of nodes in the meridional structure of the wave. This is covered in more detail by Ahlquist (1982). Fig. 1, reproduced from Ahlquist (1982), portrays the normal modes' latitudinal structure, $(U(\varphi),V(\varphi),H(\varphi))$, i.e. the Hough functions. Only the curves in dashed lines are of interest here. The curves in solid lines in fig. 1 demonstrate that the largest scale normal modes are not altered significantly when a resting basic state is replaced by a moving basic state. We shall follow standard notation and let (m,l) denote the wave whose zonal wavenumber is "m" and whose meridional wavenumber is "l."

A word is in order on interpreting nondimensional amplitudes A(p,t) and B(p,t). The Hough functions (U,V,H) graphed in fig. 1 of Ahlquist (1982) are of order 1, so a mode's velocity (u,v) is of order $(gh_e)^{1/2} = 300$ m/s times its nondimensional amplitude and its height perturbation is of order $h_e = 10^4$ m times the nondimensional amplitude. Thus, a nondimensionl amplitude of 0.01 corresponds to a velocity of order 3 m/s and a height perturbation of 100 m.

3.2 Forecast error as seen in daily plots of analyses and forecasts

In this study, we examined waves indexed by zonal wavenumbers m = 1, 2, 3, 4 and meridional indices l = 1, 2, 3, 4. The five pages of fig. 2 are plots of the amplitude and phase of the ordered pair (A(p,t), B(p,t)) for each of the 16 modes in question for 17-21 February 1979. The five pages of fig. 3 are analogous graphs for 10-14 June 1979. The analyzed values are drawn with solid lines and the forecasted values with dashed lines.

We begin by noting usual characteristics of the analyzed amplitude and phase and follow that by discussing characteristics of the forecast errors. It is evident that modal amplitude usually grows with height in the troposphere and decreases with height in the stratosphere. Phase usually shifts westward with increasing height. These are not precisely the features of theoretical barotropic normal modes in an isothermal atmosphere, which grow with height at all levels as a Lamb wave (velocity and height perturbations growing as p^{-R/c_p}) and having no phase tilt with height). Sudden phase shifts with height are visible for the (3,1) mode during 17-18 February 1979, for the (3,3) mode during 17-20 February 1979, and for the (1,1) mode during 10 June 1979. This latter behavior characterizes internal, baroclinic modes.

Forecast errors will be discussed as a function of mode, height, and time. In the February forecast, the strength of the zonal wavenumber 4 modes was consistently too weak (by a factor of two at 48 hours and later), but zonal wavenumber 4 was handled better in the June forecast, except for the (4,4) mode. As

height increases, errors tend to increase, but not strongly, and there are cases where the lower tropospheric error is as large as or larger than the upper tropospheric error. The large error in amplitude at 10 kPa (100 mb) on 19 February 1979 is presumably due to some error in scaling when AFGL saved the forecast output because the amplitude error appears at only one level and the phase appears consistent at all levels. As forecast time progressed, errors grew. We examine error as a function of time in more detail in the next subsection.

3.3 Forecast error at 50 kPa as a function of time

Next we look at the time evolution of the forecast at 50 kPA (500 mb), near the level of nondivergence, where our barotropic perspective is most meaningful. The time evolution of the analyses and forecasts for 50 kPa are plotted in fig. 4 for 17-21 February 1979 and 10-14 June 1979. For each mode, A(50kPa,t) is plotted on the abscissa and B(50kPa,t) is plotted on the ordinate. The analyzed values are drawn with solid lines and the forecasted values with dashed lines. A black dot marks values at the beginning of the forecast, i.e. a persistence forecast.

We note that the analyses for zonal wavenumbers 1 through 3 drift westward (clockwise) with increasing time, as we would expect for Rossby waves. Zonal wavenumber 4 drifts eastward (counterclockwise), more characteristic of synoptic-scale activity.

No consistency of error as a function of mode index is

apparent. The forecasts for the (3,3) mode during February and the zonal wavenumber 2 modes during June are among the most accurate. The forecasts for the (3,3), (4,1), and (4,2) modes during June were among the worst. Even so, all the forecasts are better than persistence throughout the two 4-day periods.

3.4 Statistical forecasting results

Since no forecast can be expected to be absolutely perfect, a question arises as to how much skill the forecast has relative to a fairly unsophisticated forecast. One of the most common unsophisticated forecasts is to assume that weather will persist unchanged "in situ." We felt that "in situ" persistence would be an oversimplified forecast for our investigation because wave propagation is expected. [See Ahlquist (1982, 1985) and Lindzen et al. (1984)].

As explained in section 2, we chose to use a statistical forecasting technique which allows the waves to propagate and decay. In order to estimate the parameters for the statistical forecasting model, we needed analyses prior to each forecast period. Such analyses were not part of the data obtained from AFGL, so we chose to forecast for Ahlquist's (1982) observed time series (A(t), B(t)) for the various normal modes. Ahlquist's time series extend twice daily from July 1976 through September 1979 and were determined by fitting a Lamb wave structure to the NMC analyses at 85, 50, and 20 kPa (850, 500, and 200 mb) and then projecting onto normal mode strucutres. Thus, they differ

somewhat from the projections graphed in fig. 4, which are based only on 50 kPa analyses. Ahlquist's time series are scaled for 100 kPa and should be multiplied by $(100 \text{ kPa} / p)^{R/c_p}$ if applied at some other pressure "p." For p = 50 kPa, the time series should be multiplied by 1.22.

In detail, model coefficients, i.e. the average and the AR coefficients, were estimated using the 30 days prior to each forecast period. This was done using a program in file HFCAST, which called a subroutine in file MEMCO. The ordered pair (A(t), B(t)) was treated as a complex number for this forecasting.

The resulting statistical forecasts for the February and June periods are shown in fig. 5. The statistical forecasts are better than persistence, which is indicated by the black dot in each frame. The forecasts for the (2,2), (3,2), (3,4), and (4,2) modes are fairly good for the February period, and forecasts for the (2,1) and (3,1) modes are good for the June period. The AFGL forecasts are better than the statistical forecasts overall, but AFGL's multi-level spectral model is considerably more complicated than the statistical method. On the other hand, planetary scales should be the easiest spatial scales to predict on time scales of several days, so a simple method could be expected to perform moderately well.

4. Summary and conclusions

The planetary-scale wind and height field have been shown to exhibit forecast errors which, at planetary scales, are unbiased with respect to mode or height. The forecasts based on the spectral model are generally better than the statistical forecasts. Given the lack of bias in the spectral model's errors at planetary scales, it is difficult to suggest what could be done to improve the forecasts. At least it is clear that the Air Force weather prediction model suffers from none of the problems uncovered by Daley et al. (1981).

Acknowledgements

This research was sponsored by an Air Force Research Initiation Program grant, administered by Universal Energy Systems, Inc. Computing facilities were provided by Florida State University (FSU) and the FSU Department of Meteorology.

Referenoes

Ahlquist, Jon E., 1982: Normal mode global Rossby waves: theory and observations. J. Atmos. Sci., 39, 193-202.

Ahlquist, Jon E., 1985: Climatology of normal mode Rossby waves. J. Atmos. Sci., 42, 2059-2068.

Box, G.E.P. and G.M. Jenkins, 1976: Time Series Analysis: Forecasting and Control, revised edition. Holden-Day, San Francisco.

Brenner, S., C.-H. Yang, K. Mitchell, 1984: The AFGL Global Spectral Model: Expanded Resolution Baseline Version. AFGL Tech. Note 84-0308.

Daley, R., J. Tribbia, and D. Williamson, 1981: The excitation of large-scale free Rossby waves in numerical prediction. Mon. Wea. Rev., 109, 1836-1861.

Lindzen, R.S., D.M. Straus, and B. Katz, 1984: An observational study of large-scale atmospheric Rossby waves during FGGE. J. Atmos. Sci., 42, 2059-2068.

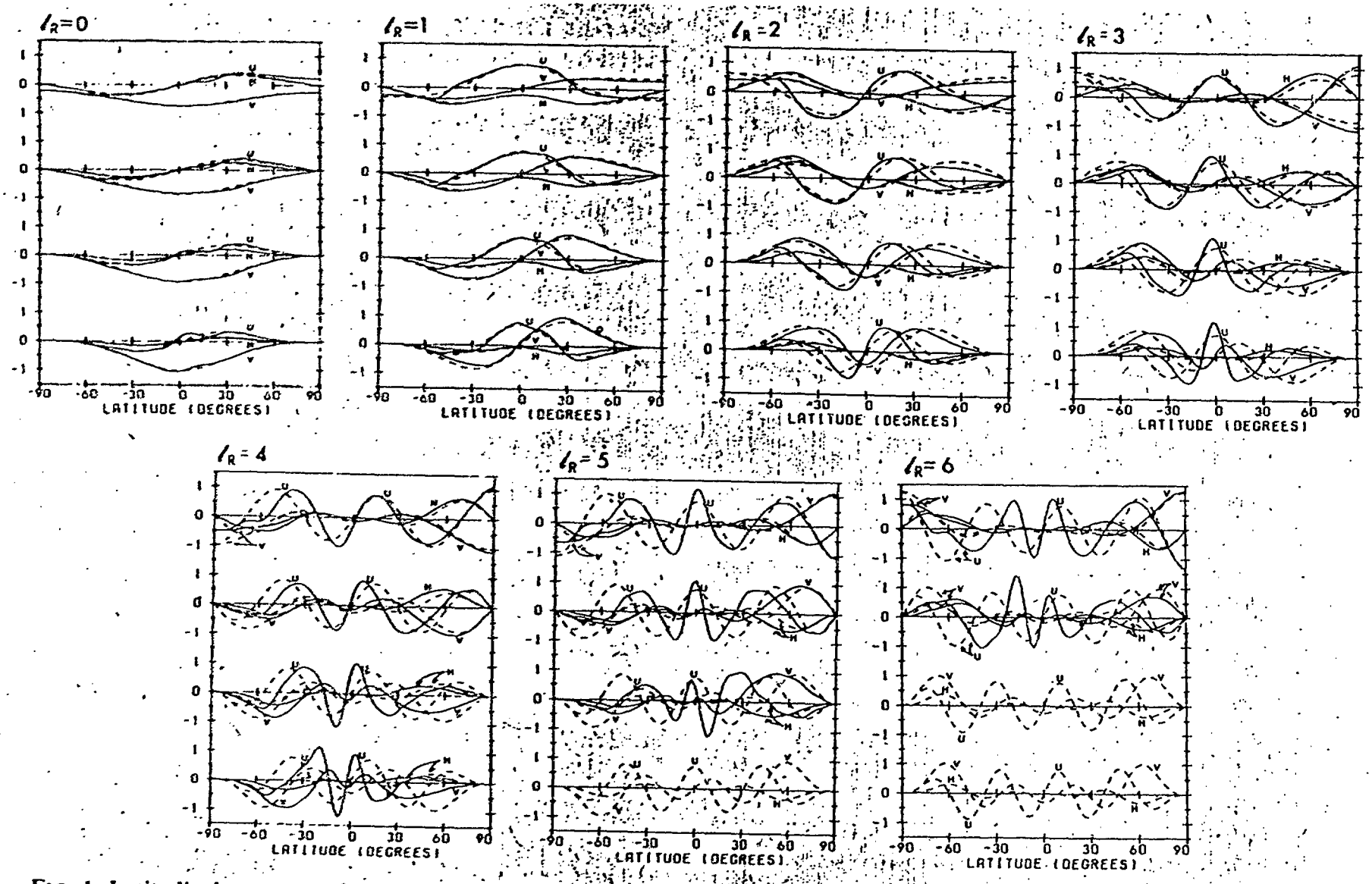


FIG. 1. Latitudinal structure of normal modes with and without the zonal wind effect. Dashed lines represent the structure in the case of no zonal flow; these are Hough functions. Solid lines represent the structure in the case of Kasahara's zonally averaged 500 mb DJF profile. The block of four graphs for each meridional index, l_R, is for zonal wavenumber $s = 1$ (at the top of each set) through $s = 4$ (at the bottom).

1200 UTC, 17 FEBRUARY 1979

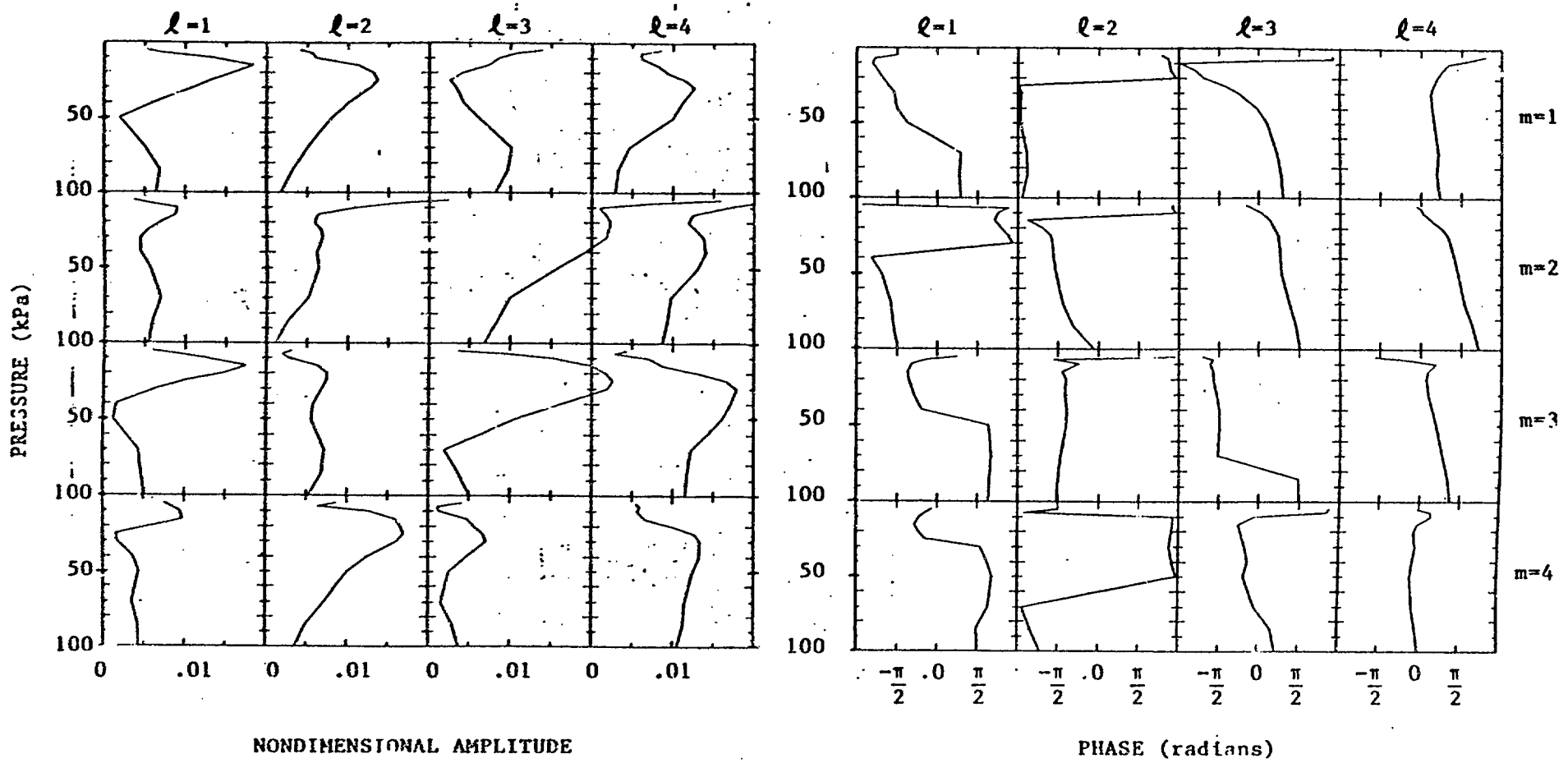


Fig. 2a: Nondimensional amplitude and phase of projections of NMC analyses onto normal mode structures. "m" denotes the zonal wavenumber of the normal mode, and "l" denotes the mode's meridional index.

1200 UTC, 18 FEBRUARY 1979

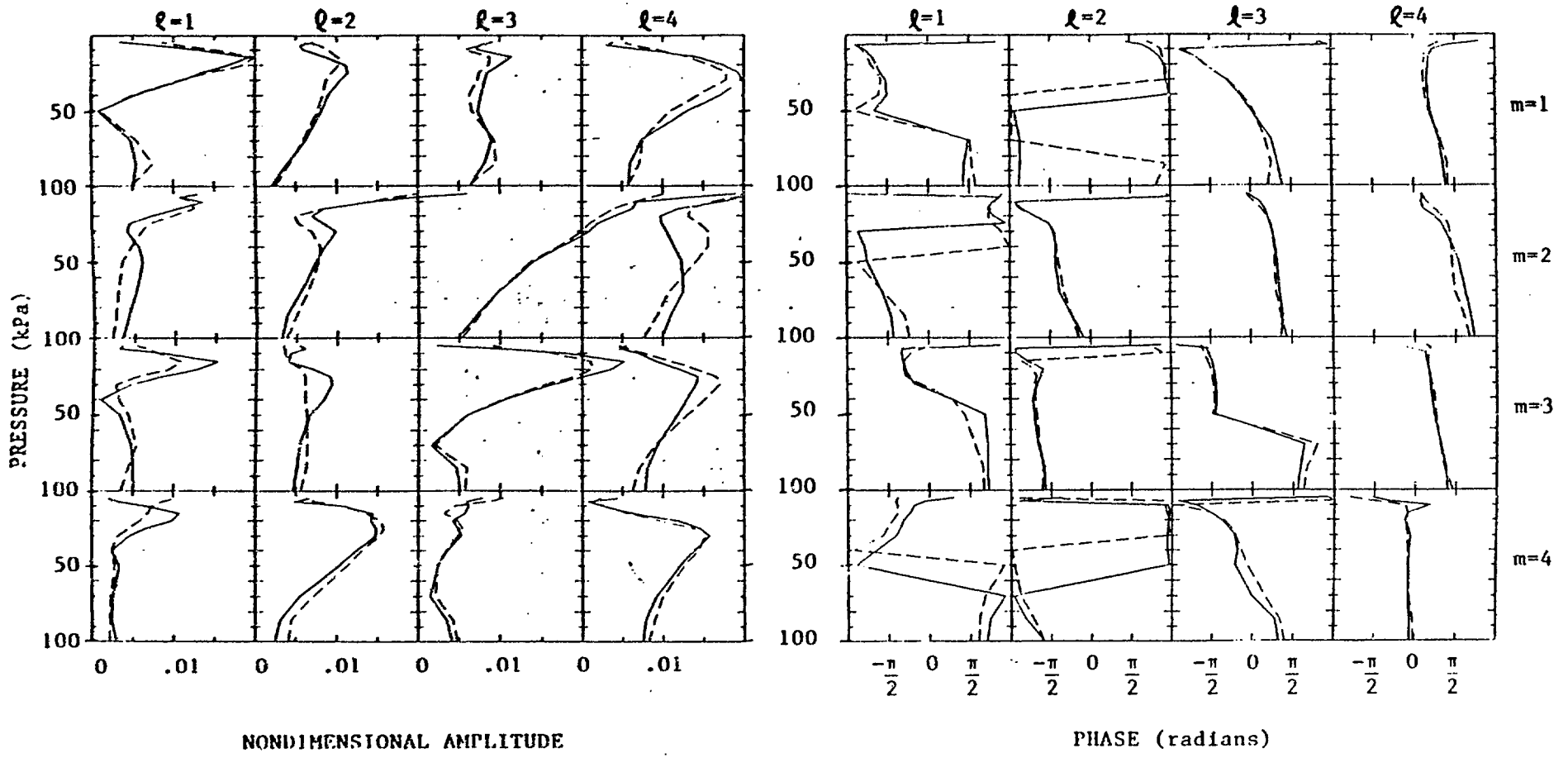


Fig. 2b: Nondimensional amplitude and phase of projections of NMC analyses onto normal mode structures. "m" denotes the zonal wavenumber of the normal mode, and "l" denotes the mode's meridional index. The amplitude and phase of projections of AFGL forecasts for the same time are drawn as dashed lines.

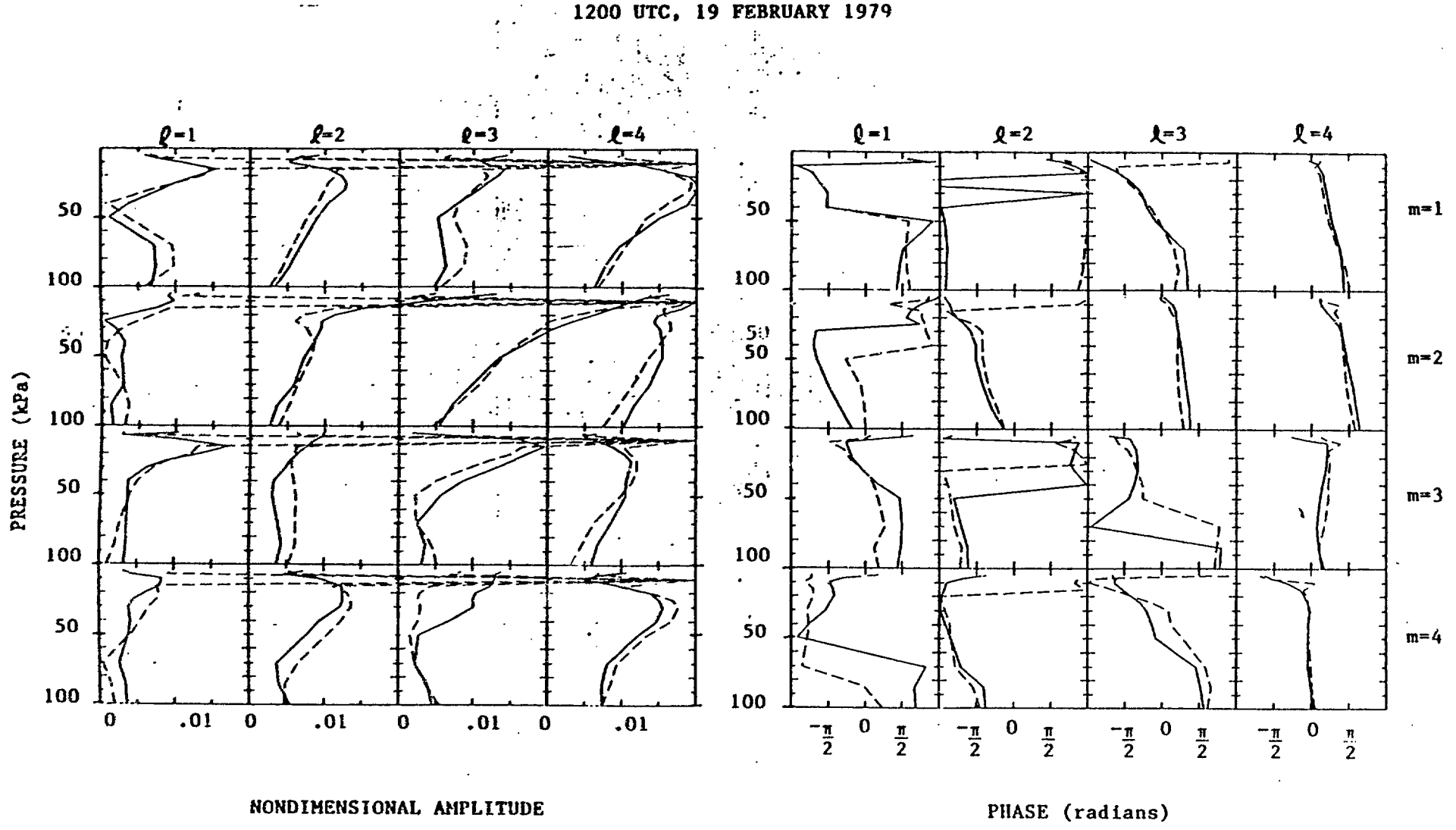
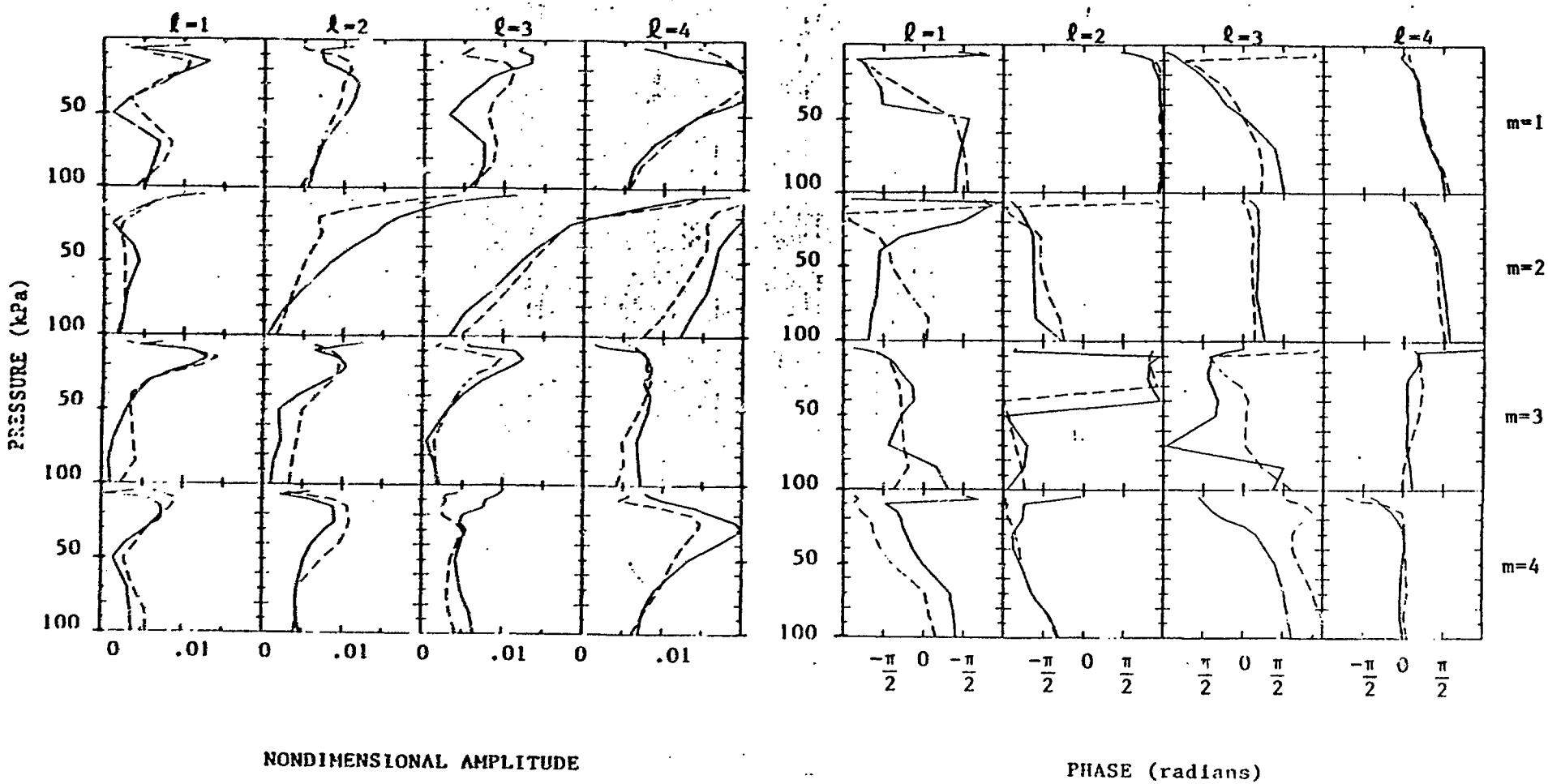


Fig. 2c: Nondimensional amplitude and phase of projections of NMC analyses onto normal mode structures. "m" denotes the zonal wavenumber of the normal mode, and "l" denotes the mode's meridional index. The amplitude and phase of projections of AFGL forecasts for the same time are drawn as dashed lines.

1200 UTC, 20 FEBRUARY 1979



NONDIMENSIONAL AMPLITUDE

PHASE (radians)

Fig. 2d: Nondimensional amplitude and phase of projections of NMC analyses onto normal mode structures. "m" denotes the zonal wavenumber of the normal mode, and "l" denotes the mode's meridional index. The amplitude and phase of projections of AFGL forecasts for the same time are drawn as dashed lines.

1200 UTC, 21 FEBRUARY 1979

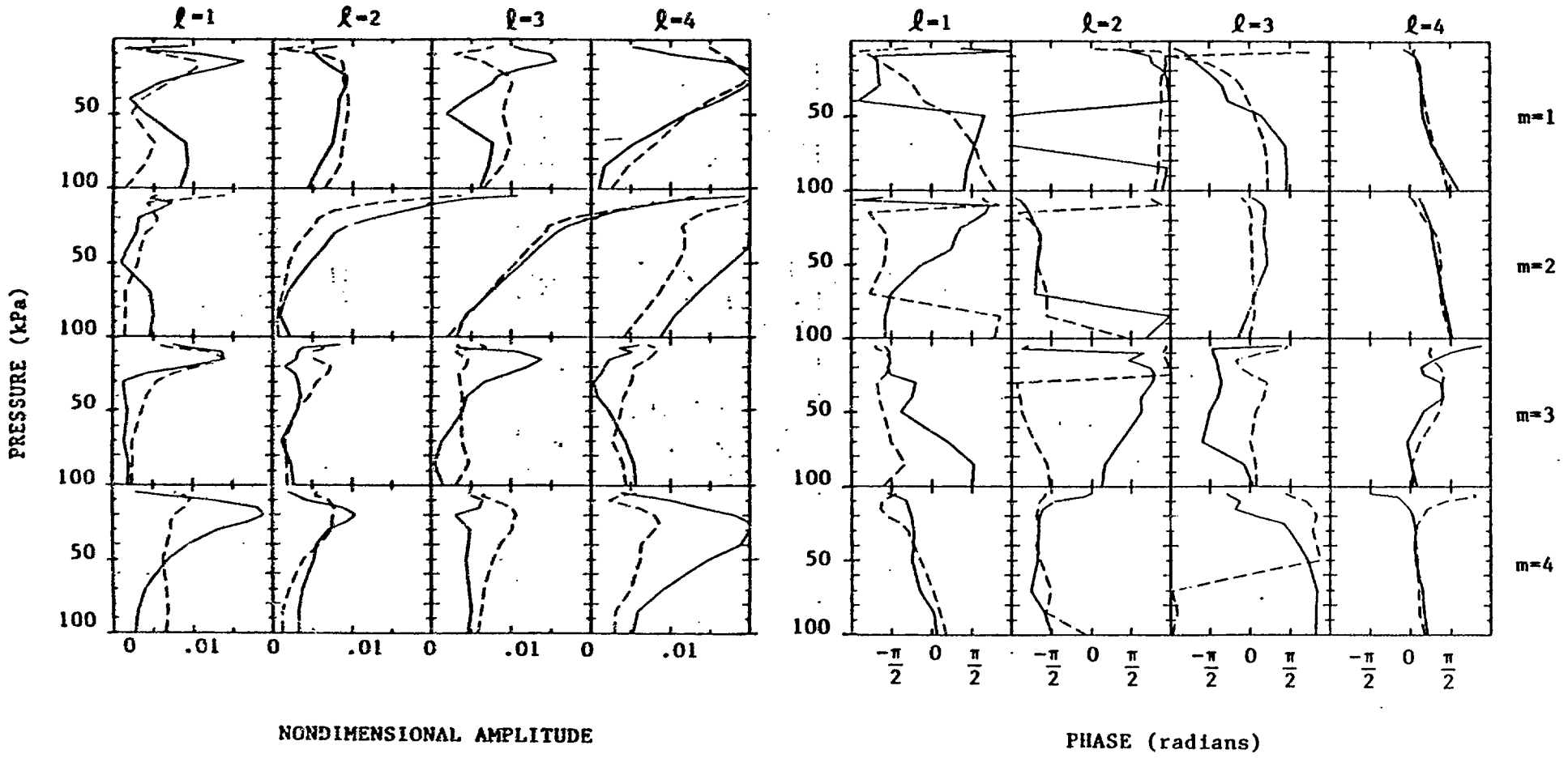


Fig. 2e: Nondimensional amplitude and phase of projections of NMC analyses onto normal mode structures. "m" denotes the zonal wavenumber of the normal mode, and "l" denotes the mode's meridional index. The amplitude and phase of projections of AFGL forecasts for the same time are drawn as dashed lines.

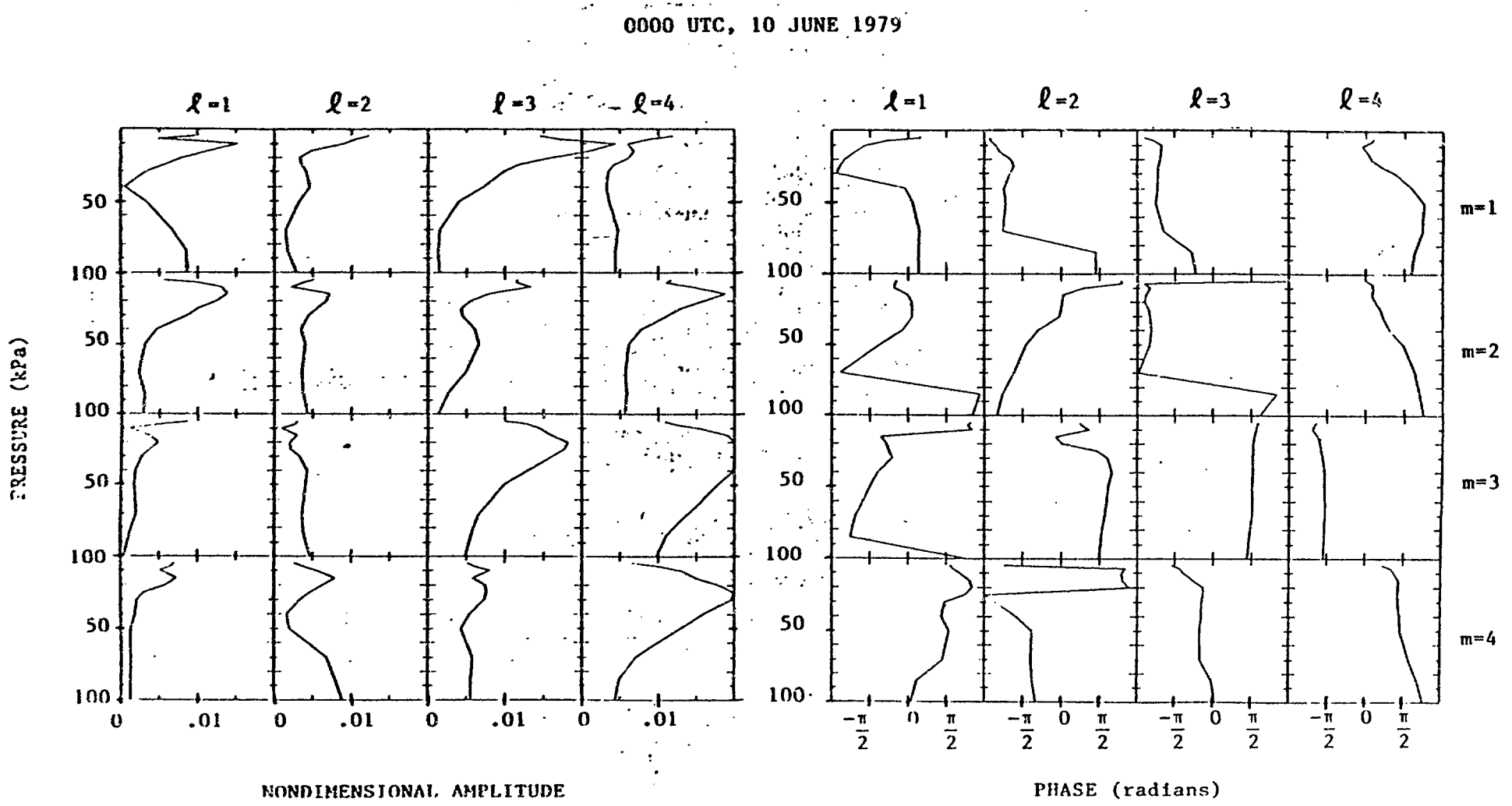


Fig. 3a: Nondimensional amplitude and phase of projections of NMC analyses onto normal mode structures. "m" denotes the zonal wavenumber of the normal mode, and "l" denotes the mode's meridional index.

0000 UTC, 11 JUNE 1979

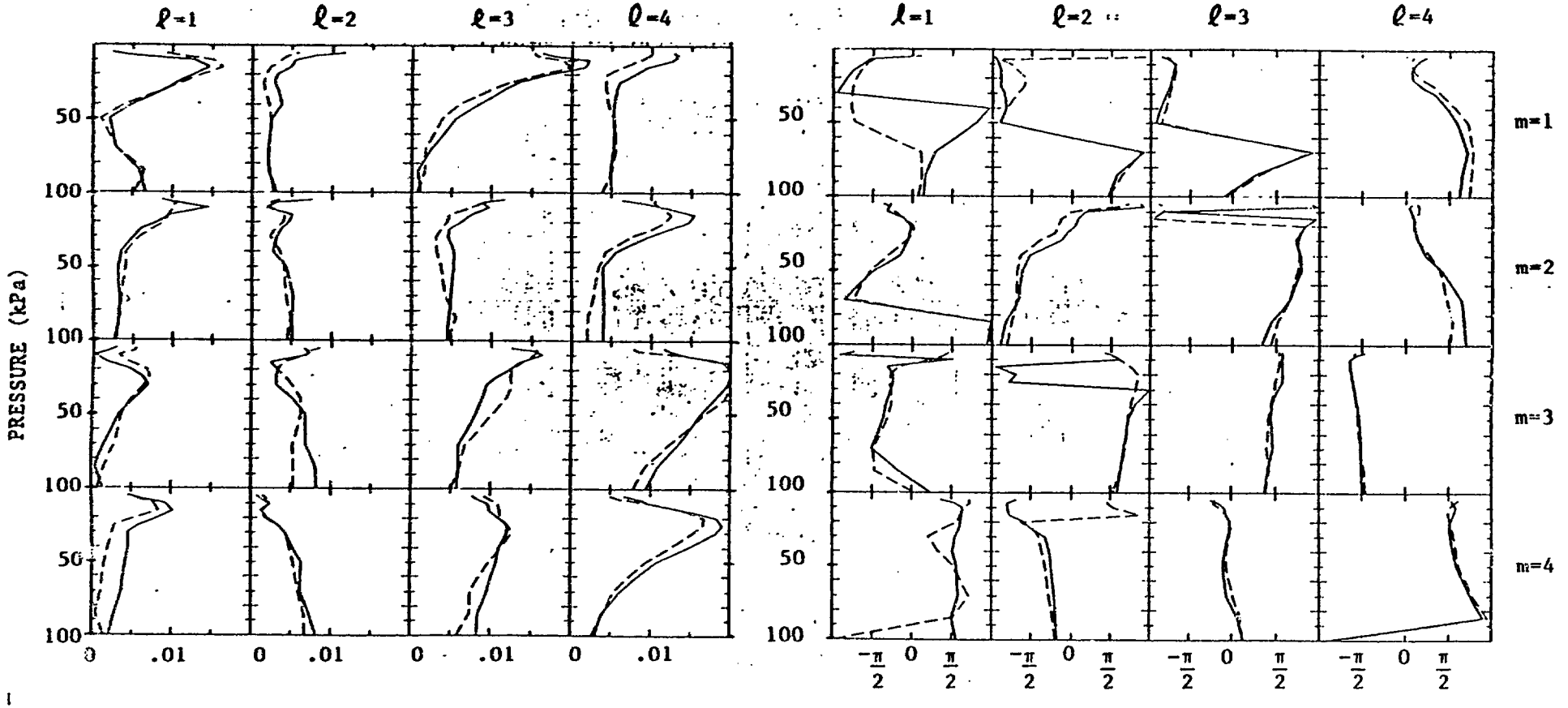


Fig. 3b: Nondimensional amplitude and phase of projections of NMC analyses onto normal mode structures. "m" denotes the zonal wavenumber of the normal mode, and "l" denotes the mode's meridional index. The amplitude and phase of projections of AFGL forecasts for the same time are drawn as dashed lines.

0000 UTC, 12 JUNE 1979

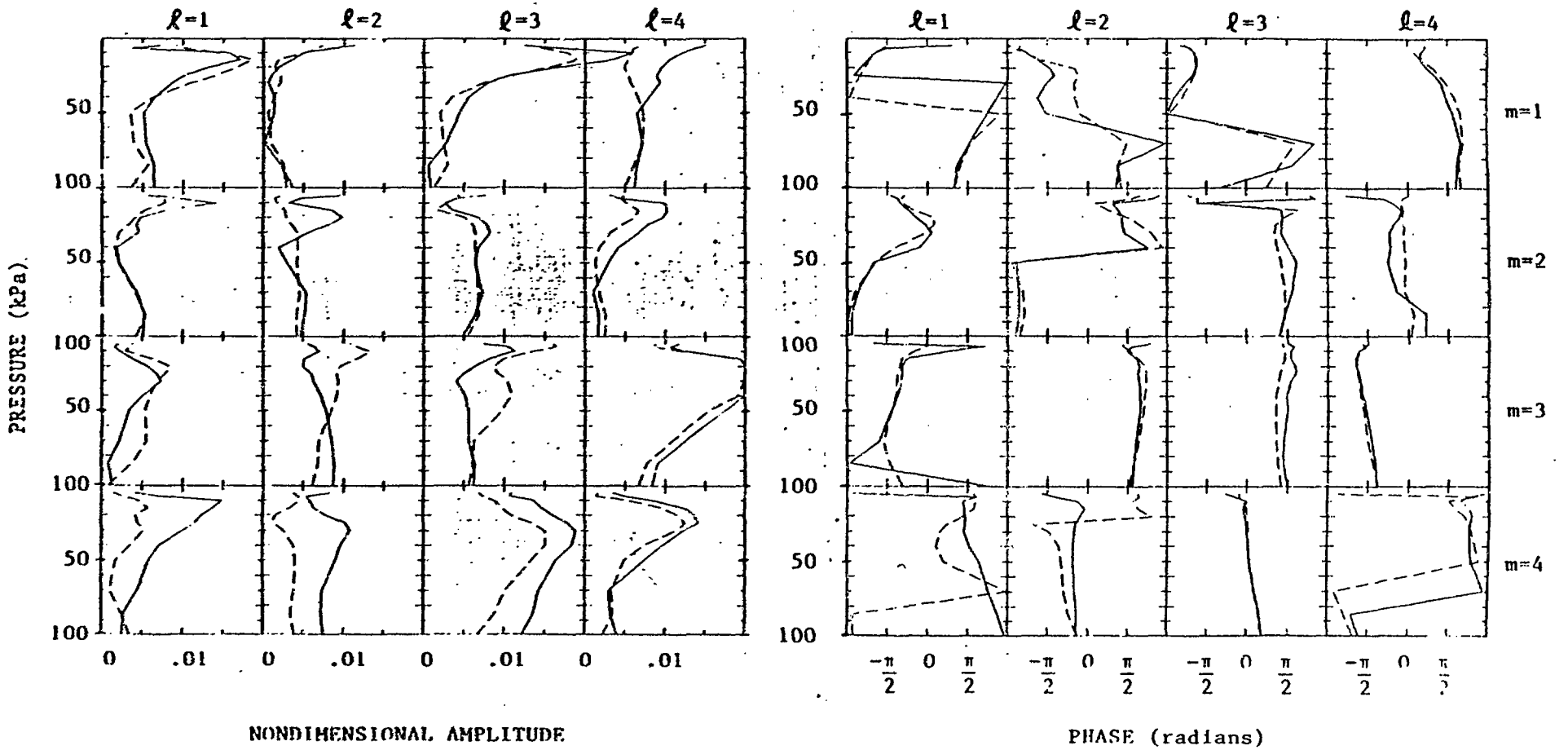


Fig. 3c: Nondimensional amplitude and phase of projections of NMC analyses onto normal mode structures. "m" denotes the zonal wavenumber of the normal mode, and "l" denotes the mode's meridional index. The amplitude and phase of projections of AFGL forecasts for the same time are drawn as dashed lines.

0000 UTC, 13 JUNE 1979

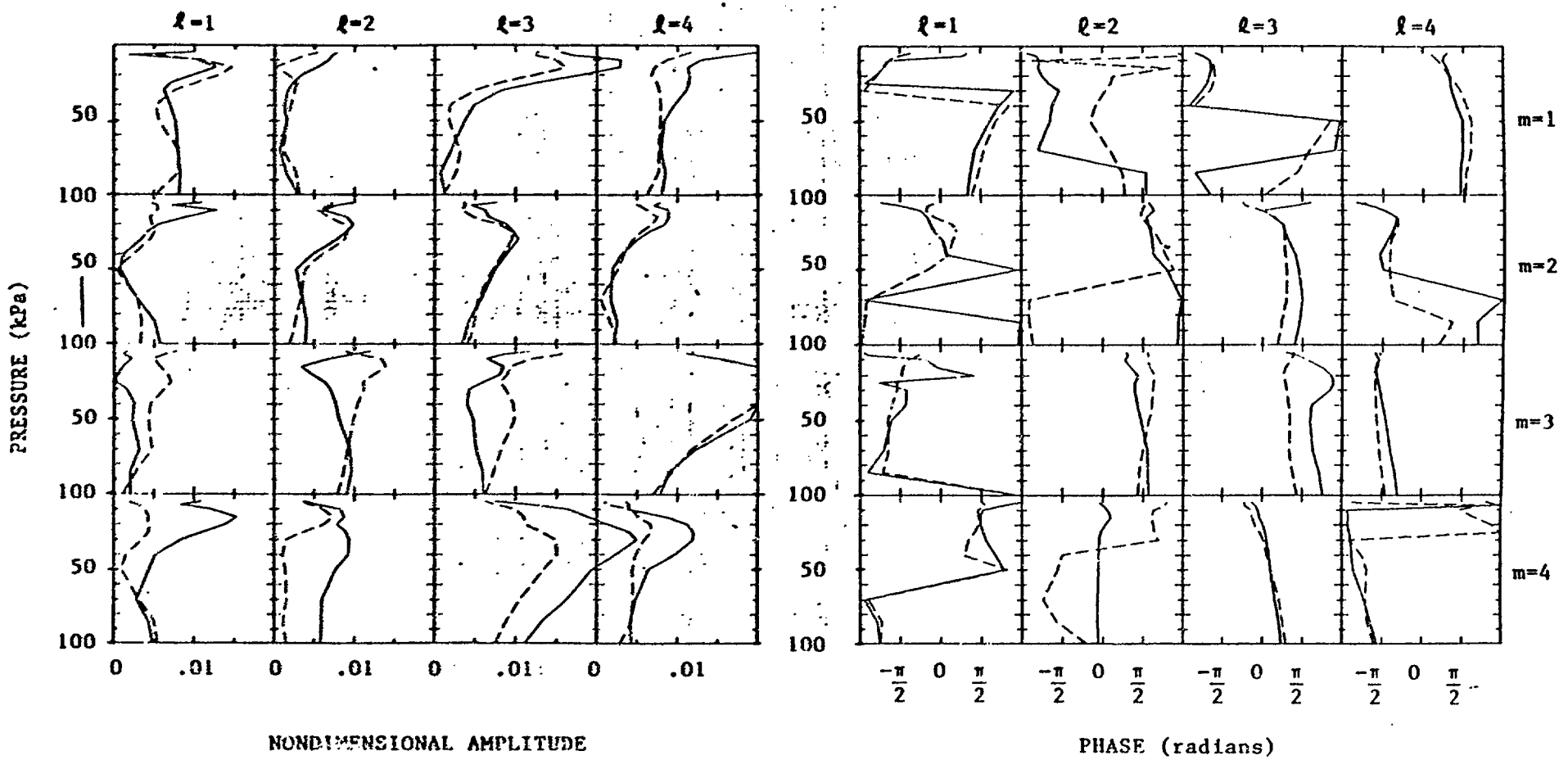


Fig. 3: Nondimensional amplitude and phase of projections of NMC analyses onto normal mode structures. "m" denotes the zonal wavenumber of the normal mode, and "l" denotes the mode's meridional index. The amplitude and phase of projections of AFGL forecasts for the same time are drawn as dashed lines.

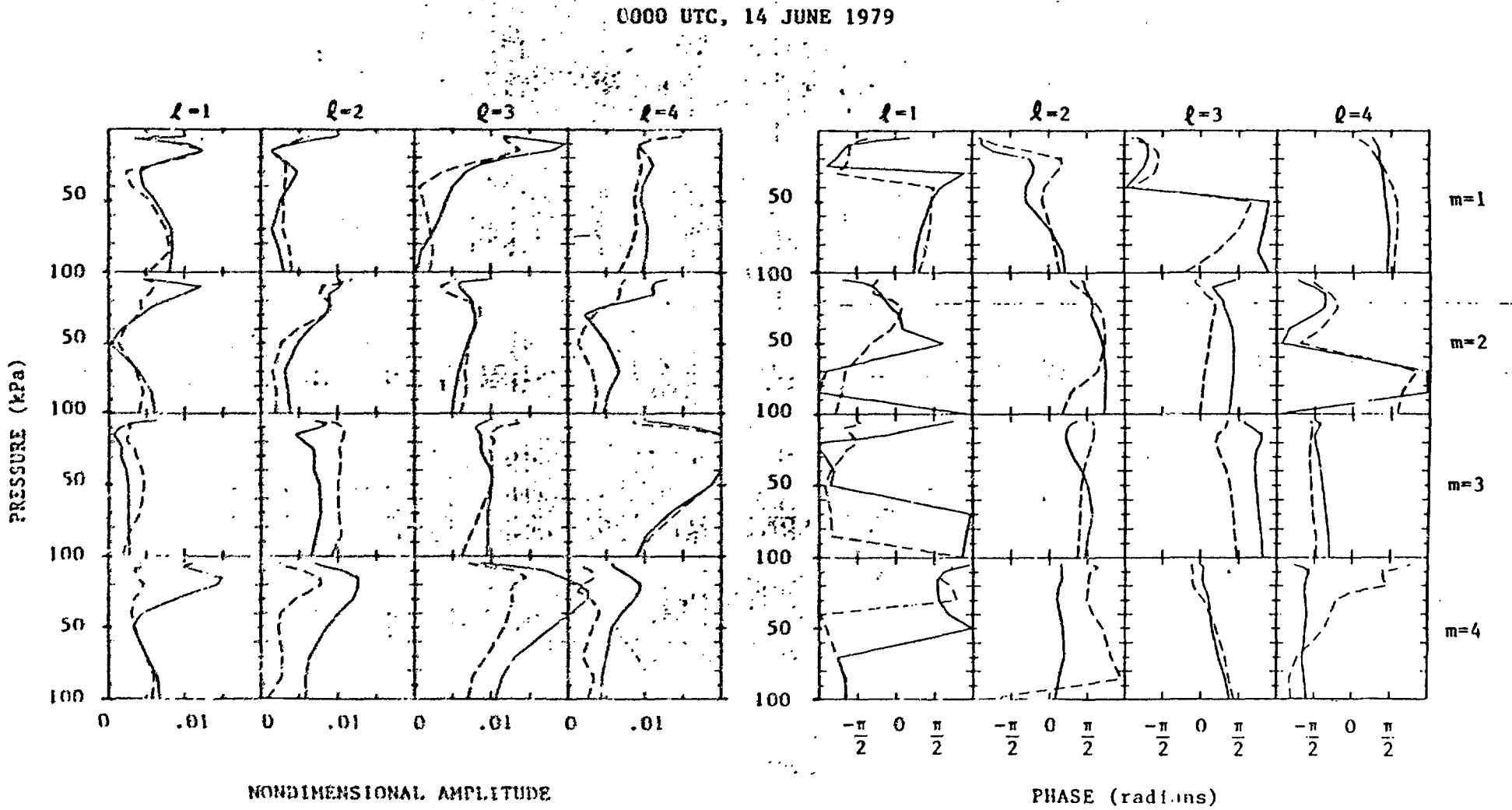
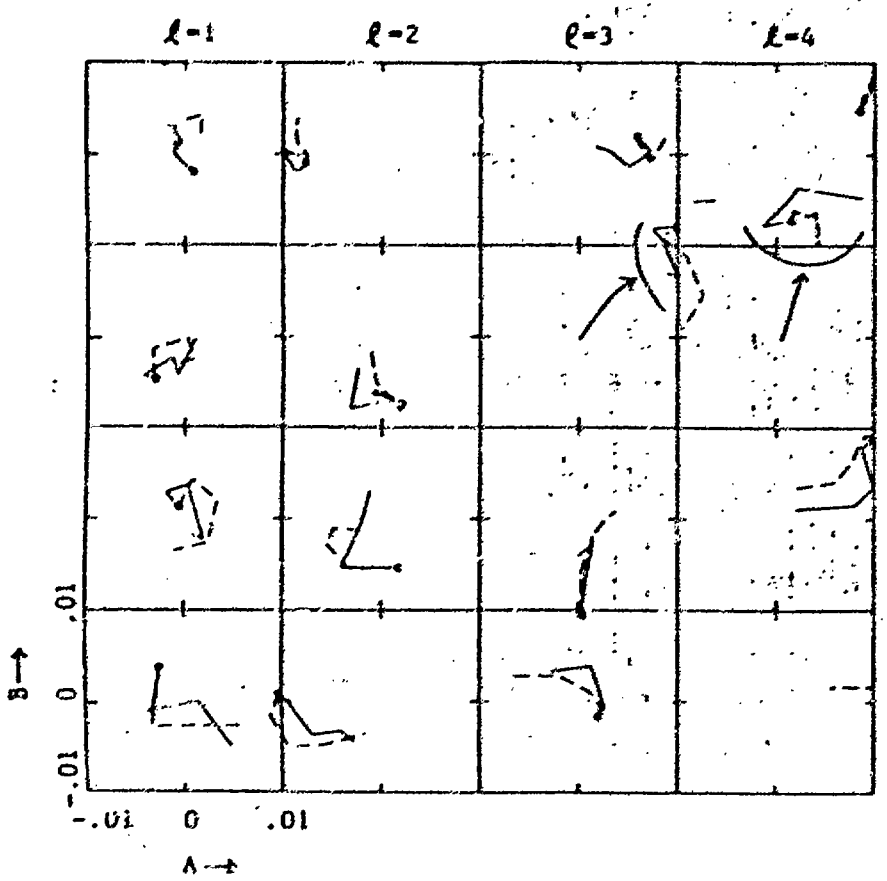


Fig. 3c: Nondimensional amplitude and phase of projections of NMC analyses onto normal mode structures. "m" denotes the zonal wavenumber of the normal mode, and "l" denotes the mode's meridional index. The amplitude and phase of projections of AFGL forecasts for the same time are drawn as dashed lines.

AFGL SPECTRAL FORECASTS

17 - 21 FEBRUARY 1979



10 - 14 JUNE 1979

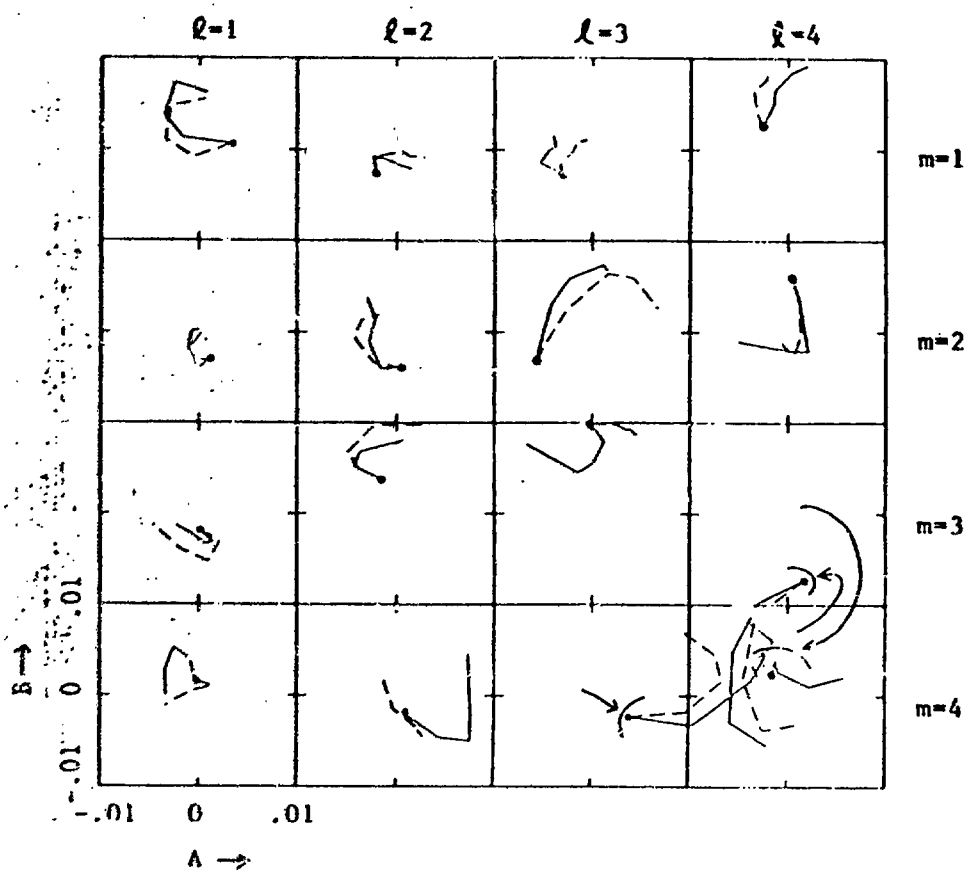


Fig. 4: Time evolution of (A(50kPa,t), B(50kPa,t)) for each projection mode. Each curve drawn with solid lines represents the evolution of the mode at 50 kPA in the NMC analyses, and each dashed curve represents the evolution of the mode at 50 kPa in the AFGL spectral forecasts. The labeling which is shown for the (m,ℓ) = (4,1) modes applies to each of the graphs. The solid black dot at the vertex of each solid and dashed line pair marks (A,B) at the starting time of the forecast and hence denotes the value for a persistence forecast. In cases where the plot for a mode trespasses into the plotting region for another mode, arrows have been drawn. An analysis and forecast pair marked by an arrow should be interpreted relative to the center of the box in which the tail of an arrow lies.

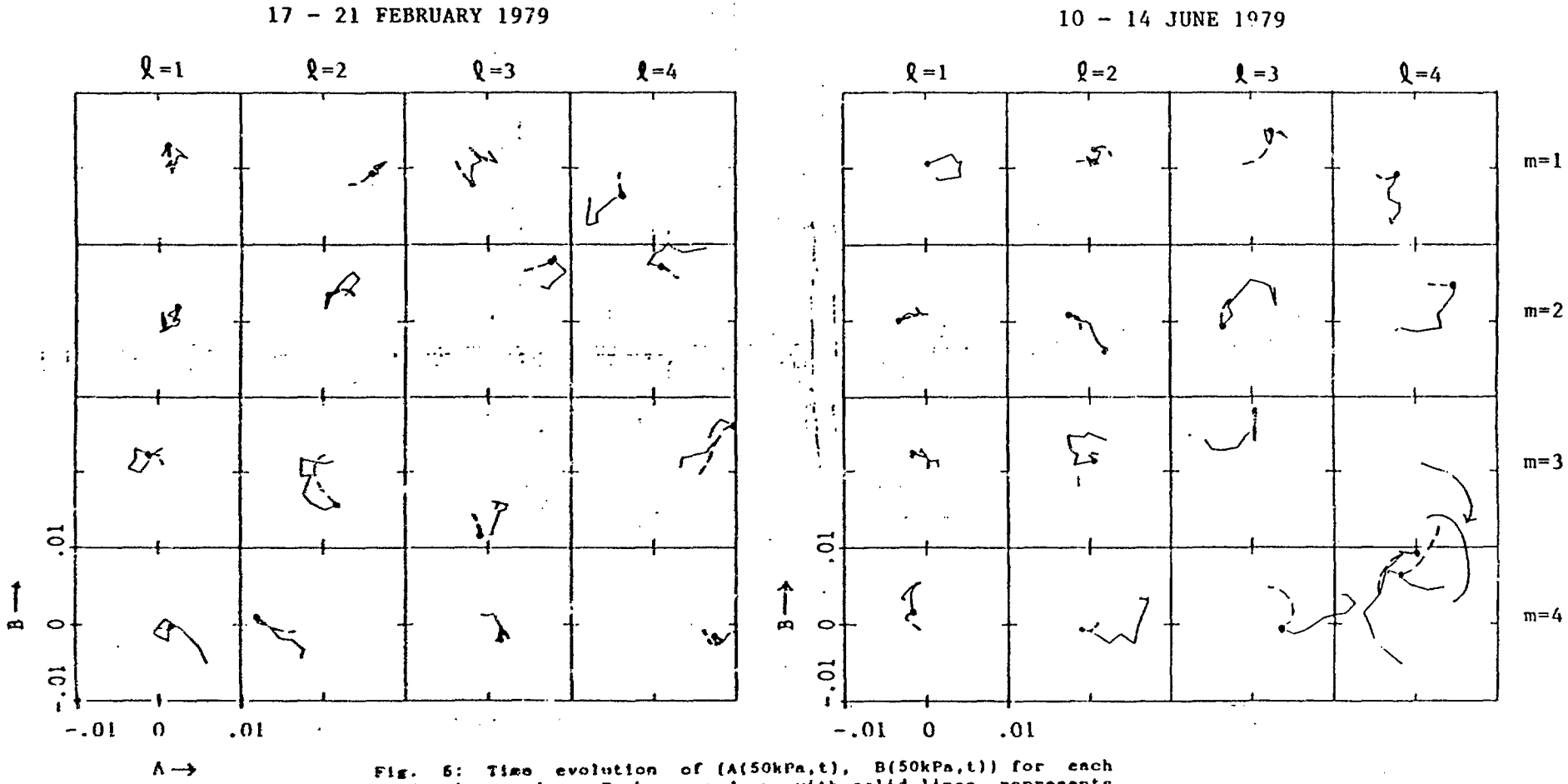


Fig. 6: Time evolution of (A(50kPa,t), B(50kPa,t)) for each projection mode. Each curve drawn with solid lines represents the evolution of the mode in Ahlquist's (1982) Rossby mode time series, and each dashed curve represents the evolution of the mode in statistical forecasts made using a second order autoregressive model. The labeling which is shown for the (m,l) = (4,1) modes applies to each of the graphs. The solid black dot at the vertex of each solid and dashed line pair marks (A,B) at the starting time of the forecast and hence denotes the value for a persistence forecast. In cases where the plot for a mode trespasses into the plotting region for another mode, arrows have been drawn. An analysis and forecast pair marked by an arrow should be interpreted relative to the center of the box in which the tail of an arrow lies.

APPENDIX A3.2 A RECURSIVE METHOD FOR CALCULATING ESTIMATES OF AUTOREGRESSIVE PARAMETERS

We now show how Yule–Walker estimates for the parameters of an AR($p+1$) process may be obtained when the estimates for an AR(p) process, fitted to the same time series, are known. This recursive method of calculation can be used to approximate the partial autocorrelation function, as described in Section 3.2.6.

To illustrate the recursion, consider the equations (3.2.34). Yule–Walker estimates are obtained for $k = 2, 3$, from

$$\begin{aligned} r_2 &= \hat{\phi}_{21} r_1 + \hat{\phi}_{22} \\ r_1 &= \hat{\phi}_{21} + \hat{\phi}_{22} r_1 \end{aligned} \tag{A3.2.1}$$

and

$$\begin{aligned} r_3 &= \hat{\phi}_{31} r_2 + \hat{\phi}_{32} r_1 + \hat{\phi}_{33} \\ r_2 &= \hat{\phi}_{31} r_1 + \hat{\phi}_{32} + \hat{\phi}_{33} r_1 \\ r_1 &= \hat{\phi}_{31} + \hat{\phi}_{32} r_1 + \hat{\phi}_{33} r_2 \end{aligned} \tag{A3.2.2}$$

where $r_k \equiv \frac{1}{N} \sum_{t=1+k}^{N} (x_t - \bar{x})(x_{t-k} - \bar{x})^*$, $(\,)^*$ = complex conjugate

The coefficients $\hat{\phi}_{31}$ and $\hat{\phi}_{32}$ may be expressed in terms of $\hat{\phi}_{33}$, using the last two equations of (A3.2.2). The solution may be written in matrix form

$$\begin{pmatrix} \hat{\phi}_{31} \\ \hat{\phi}_{32} \end{pmatrix} = \mathbf{R}_2^{-1} \begin{pmatrix} r_2 - \hat{\phi}_{33} r_1 \\ r_1 - \hat{\phi}_{33} r_2 \end{pmatrix} \tag{A3.2.3}$$

where

$$\mathbf{R}_2 = \begin{bmatrix} r_1 & 1 \\ 1 & r_1 \end{bmatrix}$$

Now (A3.2.3) may be rewritten

$$\begin{bmatrix} \hat{\phi}_{31} \\ \hat{\phi}_{32} \end{bmatrix} = \mathbf{R}_2^{-1} \begin{bmatrix} r_2 \\ r_1 \end{bmatrix} - \hat{\phi}_{33} \mathbf{R}_2^{-1} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \tag{A3.2.4}$$

Using the fact that (A3.2.1) may be rewritten

$$\begin{bmatrix} \hat{\phi}_{21} \\ \hat{\phi}_{22} \end{bmatrix} = \mathbf{R}_2^{-1} \begin{bmatrix} r_2 \\ r_1 \end{bmatrix}$$

it follows that (A3.2.4) becomes

$$\begin{bmatrix} \hat{\phi}_{31} \\ \hat{\phi}_{32} \end{bmatrix} = \begin{bmatrix} \hat{\phi}_{21} \\ \hat{\phi}_{22} \end{bmatrix} - \hat{\phi}_{33} \begin{bmatrix} \hat{\phi}_{22}^* \\ \hat{\phi}_{21}^* \end{bmatrix}$$

that is

$$\begin{aligned} \hat{\phi}_{31} &= \hat{\phi}_{21} - \hat{\phi}_{33} \hat{\phi}_{22}^* \\ \hat{\phi}_{32} &= \hat{\phi}_{22} - \hat{\phi}_{33} \hat{\phi}_{21}^* \end{aligned} \tag{A3.2.5}$$

To complete the calculation of $\hat{\phi}_{31}$ and $\hat{\phi}_{32}$, we need an expression for $\hat{\phi}_{33}$. On substituting (A3.2.5) in the first of the equations (A3.2.2), we obtain

$$\hat{\phi}_{33} = \frac{r_3 - \hat{\phi}_{21} r_2 - \hat{\phi}_{22} r_1}{1 - \hat{\phi}_{21}^* r_1 - \hat{\phi}_{22}^* r_2} \tag{A3.2.6}$$

Thus, the partial autocorrelation $\hat{\phi}_{33}$, is first calculated from $\hat{\phi}_{21}$ and $\hat{\phi}_{22}$, using (A3.2.6), and then the other two coefficients, $\hat{\phi}_{31}$ and $\hat{\phi}_{32}$, may be obtained from (A3.2.5).

In general, the recursive formulae, which are due to Durbin [34], are

$$\hat{\phi}_{p+1,j} = \hat{\phi}_{pj} - \hat{\phi}_{p+1,p+1} \hat{\phi}_{p,p-j+1}^* \qquad j = 1, 2, \ldots, p \tag{A3.2.7}$$

$$\hat{\phi}_{p+1,p+1} = \frac{r_{p+1} - \sum_{j=1}^{p} \hat{\phi}_{pj} r_{p+1-j}}{1 - \sum_{j=1}^{p} \hat{\phi}_{pj}^* r_j} \tag{A3.2.8}$$

Reproduced From
Best Available Copy

Appendix B List of ASCII files used in this research which have been supplied to researchers in the Atmospheric Prediction Branch at AFGL. Programs which would most likely be useful to AFGL researchers for general applications are underlined.

File	Contents
AFGLFEBA	normal mode amplitudes for analyses for February
AFGLFEBF	normal mode amplitudes for forecasts for February
AFGLJUNA	normal mode amplitudes for analyses for June
AFGLJUNF	normal mode amplitudes for forecasts for June
BITMOV	subroutine written in C to repack bit strings.
CALENDAR	subroutines to convert between day, month, and year and the number of days since 1 January 1 A.D. (These routines find general use in converting between date and day number.)
GBYTES	subroutine written in C which is functionally aquivalent to NCAR's GBYTES for repacking bit strings. GBYTES is less general than BITMOV.
HEXPCOE2	expansion coefficients (30 for each of U, V, and H) for a spectral representation of Hough functions. These were computed by program HNORMODG and were reordered and rescaled by a program in HUTILITY.
HFCAST	statistical forecasting program for planetary scale waves with graphical output using NCAR graphics. The programs in this file created figs. 4 and 5 for this report.
HFUNCTNS	table of Hough functions computed by a program in HUTILITY.
HNORMODG	program to find Hough functions using a Galerkin method.
HPLOT	program to plot the amplitude and phase as a function of height of the Hough function projections of the NMC analyses and the AFGL spectral forecasts. This program created figs. 2 and 3 for this report.
HPROJCT2	program to project grids onto normal mode Rossby waves, followed by a program to test the routine.
HUTILITY	programs to: 1) reorder, adjust the sign, and normalize Hough functions. 2) compute and store a table of Hough functions. 3) graph Hough functions.

MEMCO routines to fit autoregressive models to complex-valued time series and to compute the spectrum of a complex autoregressive process.

MEMRL routines to fit autoregressive models to real-valued time series and to compute the spectrum of a real autoregressive process.

Appendix C Comments on the AFGL Research Initiation Program.

Both the summer research program and the Research Initiation Program have given me an increased perception of what meteorological research is of interest to the Air Force. I am now advising an Air Force captain (Capt. Glenn Norrie) as he works for a Ph.D. in meteorology. The general area of his research will be limits to predictability resulting from "chaos," i.e. strongly nonlinear dynamics where solutions are extremely sensitive to initial conditions.

One of the spin-off benefits of my summer at AFGL is that one of my other Ph.D. students has been able to use much of the software which I wrote during my summer at AFGL for his numerical modeling research. This has contributed significantly to progress on his degree.

My own work for the Air Force proceeded much faster during the summer session at AFGL than during the follow-up work at Florida State University. Two of the more important reasons were:

1. During the summer at AFGL, learning the computer system at AFGL went quickly, and AFGL people were good about answering questions, both computational and scientific. At Florida State University, I am much more familiar with the computer system, but that did not outweigh the personal interaction factor which was so helpful at AFGL.

2. It was nontrivial to read the AFGL data tapes at Florida State University because our Cyber computer stores files to tape in a different format than does the AFGL Cyber computer. I expect problems reading tapes written by other companies' computers, but I was surprised when a problem existed between CYBER computers. For that reason, I shipped the software listed in appendix B to AFGL researchers on IBM-compatible floppy disks.

1986 USAF-UES RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by

Universal Energy Systems, Inc.

FINAL REPORT

An EPR Study of the Role of Catalysts in the Thermal Decomposition of Nitroaromatic Compounds

Prepared by:	David R. Anderson
Academic Rank:	Assistant Professor
Department and University:	Department of Chemistry
	University of Colorado at Colorado Springs
Research Location:	The Frank J. Seiler Research Laboratory,
	NC, X, United States Air Force Academy
USAF Researcher:	Captain Jon T. Swanson
Date:	January 31, 1988
Contract No:	F49620-85-C-0013/SB5851-0360

An EPR Study of the Role of Catalysts in the Thermal Decomposition of Nitroaromatic Compounds

by David R. Anderson

ABSTRACT

The intermediate nitroxide radicals produced in the thermal decompositions of dinitrotoluenes, neat and in the presence of hexamethylbenzene, have been identified by electron paramagnetic resonance (EPR) spectroscopy. The activation energies for the reactions were also measured. These results have led to a better understanding of the mechanisms for the reactions of nitroaromatic compounds and are presented in a paper submitted to the Journal of Physical Chemistry.

Colored products are formed in mixtures of nitroaromatics and amines, in both solid mixtures and dilute solutions. The interactions that give rise to these products have been characterized spectrophotometrically as charge-transfer complex formation and acid-base reaction.

The thermal decompositions of mixtures of trinitrotoluene and electron-donor charge-transfer agents (various aromatic amines) have been examined by EPR. Also examined was picramide, a nitroaromatic compound which contains an amine group. In cases where adventitious oxidation of the amine could be minimized, the weak radical spectra which resulted still could not be interpreted unambiguously.

A program written in the ASYST scientific language to interface an IBM AT computer to the EPR instrument has been completed. The program, which handles all instrument control and data acquisition and analysis, is described and made available for distribution in a paper in preparation for submission to the Journal of Chemical Education.

I. INTRODUCTION

Reaction mechanisms for the decomposition of highly energetic organic compounds have long been of interest, from the perspectives of physical organic and theoretical chemistry, as well as from applications such as propellants and explosives for commercial and defense uses. The thermal decomposition of 2,4,6-trinitrotoluene (TNT), in particular, has been studied for over 50 years. Only recently have researchers begun to probe in detail the initial steps in the free radical decomposition mechanism. Of the plethora of methods applied to the problem, electron paramagnetic resonance (EPR) spectroscopy has been the first to allow direct observation and characterization of the radical intermediates produced in the early stages of the decomposition.[1-4]

In my Summer Faculty Research project I examined the regiochemical aspects of the reactions of dinitrotoluenes (DNTs), both neat and in the presence of hexamethylbenzene (HMB).[5] Based on the preliminary results of that study and related work[6], the following projects were initiated: (1) accurate identification of the nitroxides produced from DNTs using hyperfine coupling constants determined from an extensive series of model nitroxides[6], thereby providing a more detailed description of the decomposition mechanism, (2) EPR investigation of the reactions of picramide and of TNT in the presence of aromatic amines, (3) spectroscopic characterization of the interactions that lead to colored products in mixtures of nitroaromatics and amines, and (4) further development and publication of the EPR computer interface program.

II. OBJECTIVES OF THE RESEARCH EFFORT

The overall objective of the research project was to understand in detail the mechanisms of the thermal decompositions of nitroaromatic compounds. TNT, in particular, has received considerable attention, and recent reports from this laboratory have provided much of what is known about the initial steps in the free-radical mechanism.

The specific objectives for the present work were those outlined in the recommendations from my SFRP report and in my minigrant proposal, namely:

1. To complete the characterization of the nitroxide radical intermediates in, and the kinetic measurements of, the thermal decompositions of DNTs, both neat and in the presence of hexamethylbenzene. We previously had simulated the spectra of these radicals by using hyperfine coupling constants which were estimated from only a few known nitroxides, but recent work in this laboratory has allowed us to determine these values accurately for a series of model nitroxides[6]. Thus we could provide a reliable identification of the more complex radicals from the reactions of DNTs.
2. To examine other charge-transfer agents, namely aromatic amines, as catalysts for the thermal decomposition of nitroaromatics and to identify any radical intermediates produced in the reactions.
3. To characterize spectroscopically the colored products which result from interactions between nitroaromatics and amines. The colored products could arise from charge transfer, electron transfer, sigma-complex formation, acid-base reaction, or other interactions.
4. To complete the development of a new system for computerized EPR instrument control, data acquisition and analysis, and spectra

simulation, using and IBM AT computer and the ASYST[7] scientific programming language.

III. EXPERIMENTAL

EPR spectra were recorded on a Varian E-109 spectrometer operating at 9.14 GHz using a standard TE_{102} rectangular cavity. An IBM Instruments temperature controller was used to maintain the temperature in the cavity to ±0.5°C, as measured with a Doric Trendicator. Spectra were collected, analyzed, and simulated using an IBM PC AT computer, interfaced to the EPR spectrometer through a TECMAR Lab Master A/D converter board, and running software written locally in the ASYST programming language.

All reagents were commercially available (Aldrich); liquids were typically Gold Label purity and were used as is, and solids were carefully recrystallized prior to use. Mixtures of solids were prepared by weighing out appropriate amounts of the components and then grinding them together to a fine powder to insure homogeneity. Because of the microwave absorption associated with the electric dipole inherent in DNTs, it was necessary to use slightly flattened EPR tubes in order to decrease microwave loss. Samples were reacted with the EPR tube open to the atmosphere.

Except for differences in sample composition and decomposition temperature, all qualitative spectra for structure analysis were recorded under identical conditions: 2000 data points/100 gauss field width, 300 signals averaged, center of field sweep locked on DPPH, 100 kHz modulation, 0.5 gauss modulation amplitude, 2×10^4 gain, 0.025 s time constant, 10 mW incident microwave power, and ≈9.14 GHz microwave frequency (measured using an Eldorado frequency counter).

Decomposition rate data were taken at isothermal conditions with the reported temperature being the average of the initial and final tempera-

tures, these ... ally within 0.5°C of each other. Within one minute of their introduction into the EPR cavity, the samples reached a temperature within 1°C of the stabilized temperature. EPR spectra were taken as often as necessary to characterize the kinetics of the decomposition, typically every 4 to 8 minutes. Relative radical concentrations were determined by measuring the maximum peak intensities in the EPR derivative spectra. The linewidths of the radical spectra did not change appreciably during the course of the kinetic run.

Solvents for UV-visible analyses were spectrophotometric quality and spectra were taken using a Hewlett-Packard diode array spectrophotometer.

IV. RESULTS AND DISCUSSION

A. EPR Analysis of the Catalyzed Thermal Decompositions of Nitrotoluenes

The results of our work on the thermal decomposition of DNTs and the discussion of the mechanism for nitroxide radical formation are presented in a paper submitted to J. Phys. Chem.,[8] and I refer the reader to that manuscript in the addendum to this report.

To examine the effects of electron-donor charge-transfer agents on the thermal decomposition reactions of nitrotoluenes (good electron acceptors), we chose to study mixtures of TNT and various aromatic amines: 1-aminonaphthalene, 2-aminobiphenyl, diphenyl amine, and triphenyl amine. We found that all the amines undergo oxidation to free radicals at elevated temperatures, presumably by molecular oxygen. The radical observed from diphenyl amine, in particular, was clearly recognizable as diphenyl nitroxide. When the amines were rigorously degassed and heated under nitrogen the oxidations were considerably reduced but not entirely eliminated. In mixtures of TNT and amines the radical spectra were weak, poorly resolved, and generally uninterpretable. The spectra were clearly dissimilar to those produced in the reactions of neat TNT or TNT/HMB mixtures, but they were in some cases similar to the spectra obtained from air oxidation of the amines. Thus it is not clear whether the radicals produced in mixtures of TNT and amines are arising from (1) electron-transfer reactions between TNT and amines, (2) adventitious oxidation of the amines and subsequent reaction of the oxidized intermediates with TNT, (3) TNT oxidation of the amine, or (4) radical reactions of the TNT itself followed by reaction with the amines.

The thermal decomposition of picramide was also of interest since it

is a species in which charge-transfer occurs intramolecularly. The strongly electron-withdrawing nitro groups are able to draw considerable electron density from the electron-donating amino group. Neat samples of picramide could not be used since they decomposed at temperatures below their melting points. Diphenyl ether was found to be a good solvent in which to carry out the reaction: it is inert, has a boiling point above the temperature required for decomposition of picramide, and is sufficiently polar to dissolve picramide. Unfortunately, the spectrum that resulted was no more well-resolved than those from the TNT/amine mixtures; the line-broadened nine-line pattern with splittings of about five gauss did not provide enough information for structure identification.

The decomposition of picramide was also carried out in the presence of HMB. The structure of the resulting radical intermediate has been tentatively identified as 4-amino-3,5-dinitrophenyl pentamethylbenzyl nitroxide. This product likely results from the same mechanism as the reaction of other nitroaromatics with HMB, namely through hydrogen atom abstraction from HMB and trapping of the benzyl radical by a nitroso group. Thus we see no effect of the charge-transfer nature of picramide in this reaction.

8. Spectroscopic Analysis of the Interactions between Nitroaromatics and Amines

During the course of examining mixtures of TNT and aromatic amines, we observed that many of the mixtures produced highly colored (orange to red) solutions when warmed. With diphenyl amine, for example, the solid mixture was colorless, but when held in the hand became a dark orange liquid; when cooled to room temperature the mixture solidified back to

colorless. We thus assayed to characterize this interaction.

The various interactions between electron acceptors like nitroaromatics and electron donors like amines can be classified as:[9] (1) charge-transfer complex formation, in which a complex is formed between the electron donor-acceptor pair, (2) electron transfer, in which complete transfer of an electron results in formation of the radical anion and cation pair, (3) acid-base reaction, in which a base removes an acidic proton from the methyl group of, for example, TNT to produce the TNT anion, and (4) sigma-complex formation, in which a zwitterionic product results from nucleophilic attack on the nitroaromatic ring. Subsequent reactions of these intermediates can produce other products. These various species have been characterized by a number of spectroscopic methods. We chose UV-visible spectrophotometry as a first step in our analysis of the TNT/aromatic amine mixtures.

Mixtures of TNT and 1,3,5-trinitrobenzene (TNB) with various aromatic amines gave solutions ranging in color from yellow to red in solvents such as methanol, acetonitrile, dimethylsulfoxide, and dimethylformamide. The amines chosen for study were aniline, p-toluidine, N-methylaniline, and 1-aminonaphthalene, since these were readily available in sufficient purity to minimize long-wavelength absorption by the amines and also represented a range of pK_b values. We also examined mixtures of TNT and TNB with tetramethylguanidine (TMG), a base which is known to give both acid-base reaction and sigma-complex formation with TNT and TNB, respectively.[10] UV-visible spectra were taken in the range of 200 to 800 nm on methanol and acetonitrile solutions that were approximately 0.10 M nitroaromatic and 0.020 M amine.

We found that the mixture of TNT and TMG indeed yielded the blue TNT

anion, characterized by absorbance maxima at 370, 520, and 650 nm, and that TNB and TMG produced the sigma complex with maxima at 440 and 500 nm.[10] The spectra resulting from mixtures of TNT and TNB with all of the aromatic amines, however, did not exhibit the long-wavelength absorbance maxima but only a broadening of the π-π^* absorbance band. The increased absorbance of the tail of this band above 400 nm gives rise to the yellow-orange-red color. This spectroscopic behavior is characteristic of the charge-transfer complexes observed in other mixtures of aromatic amines and nitroaromatics.[11] Although mixtures of TNT, TNB, and the particular amines we chose for study have not been reported, the similarity of our spectroscopic results with those in the literature suggest that the colors we observe are due to charge-transfer complex formation.

C. EPR/Computer Interface Program

In conjunction with the collection of EPR data we developed a computer interface for the EPR spectrometer which provides numerous capabilities that are not available from the EPR instrument itself. The interface involves connection of the EPR instrument to an IBM AT computer through an A/D converter, and a program to carry out its functions. We chose to use the ASYST[7] programming language because of its inherent algorithms for facile acquisition and analysis of data. The program controls the EPR instrument and collects and analyses the raw data, then allows the user to manipulate the spectra in a number of ways: display, store, and plot spectra; expand, smooth, differentiate, and integrate spectra; simulate spectra from entered hyperfine coupling terms and overlay simulated and experimental spectra; and to take spectra at timed intervals, calculate signal intensities as a function of time, and then

calculate a least-squares kinetic fit to any order reaction. We have structured the program to be "user-friendly", that is, it is completely menu-driven and almost impossible for the user to enter input that would cause an error. It has greatly improved our spectra collection and manipulation capabilities; we are preparing a publication to describe the program and announce its availability for distribution in the Journal of Chemical Education.[12]

V. CONCLUSION

The goals of this project have been realized successfully. The study of the chemistry of dinitrotoluenes has been completed and submitted for publication. The interactions between aromatic nitro compounds and amines have been characterized as charge-transfer complex formation. When thermally decomposed, these mixtures undergo reactions different from those of TNT or the DNTs themselves, but the radical products could not be identified. An EPR computer interface program has been successfully completed and a publication is in preparation.

VI. RECOMMENDATIONS

With the regiochemical aspects of the mechanism for TNT and DNT decomposition having been studied, the main recommendation for continued research is to investigate other aspects of the mechanism. Specifically, since it is thought that benzyl radicals may also be trapped by nitro groups to yield aryl benzyloxy nitroxides, it is recommended that these intermediates be searched for. They are not expected to be stable at the elevated temperatures used for the decomposition studies, but should break down initially to a benzyloxy radical and a nitroso group, the benzyloxy radicals serving as precursors for the observed oxidized benzyl species: benzyl alcohol, benzaldhyde, and benzoic acid derivatives. The aryl benzyloxy nitroxides might be generated by other means, possibly by the low-temperature photolysis of nitroaromatics or of mixtures of nitroaromatics and dibenzyl mercury.

No further recommendations are suggested regarding the interactions between nitroaromatics and amines.

The EPR interface program is essentially complete in its current form, but as in all computer programs additional capabilities may be added and minor bugs corrected. It is recommended that improvements continue to be made on the program and that upgraded versions be made available.

ACKNOWLEDGEMENTS

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of my research. I am indebted to the personnel at the Frank J. Seiler Research Laboratory for their support and intellectual stimulation. I particulary wish to thank Captain Jon T. Swanson for the opportunity to participate in this research project and for his guidance throughout the program. Also greatly appreciated are Dr. Norman Heimer (U. Mississippi) for his helpful suggestions, Mr. Phillip J. Conrad (U. Colorado) for much of the experimental work, and Mr. Mark V. O'Neill (U. Colorado) for his work on the computer program.

REFERENCES

1. Guidry, R. M. and L. P. Davis, "Thermochemical Decomposition of Explosives. I. TNT Kinetic Parameters Determined from ESR Investigations," Thermochim. Acta, 32 (1979) 1-18.
2. Swanson, J. T., L. P. Davis, R. C. Dorey, and W. R. Carper, "An EPR Study of the Thermal Decomposition of Molten 2,4,6-Trinitrotoluene and its Isotopic Analogs," Magn. Reson. Chem., 24 (1986) 1-6.
3. Swanson, J. T., S. R. Bosco, C. Y. Kruit, R. D. Murphy, and W. R. Carper, "An EPR Study of the Thermal Decomposition of TNT/HMB Mixtures," NTIS Report FJSRL-TR-85-0004, August 22, 1985.
4. Goldberg, I. B., T. M. McKinney, L. F. Warren, and J. T. Swanson, "EPR Observation of Nitroxide Free Radicals during Thermal Decomposition of 2,4,6-Trinitrotoluene and Related Compounds," J. Phys. Chem., 90 (1986) 1008-1011.
5. Anderson, D. R., "An EPR Study of the Intermediate Radicals Produced in the Catalyzed and Uncatalyzed Thermal Decomposition of Dinitrotoluenes," Final Report, Universal Energy Systems, 1986.
6. Anderson, D. R., Heimer, N. E., Swanson, J. T., and Trafton, W. E., manuscript in preparation.
7. "ASYST - A Scientific System," Version 2.00, Macmillan Software Company, New York, NY, 1985.
8. Anderson, D. R., Conrad, P. J., Swanson, J. T., Trafton, W. E., McKinney, T. M., and Goldberg, I. B., "An EPR Study of the Thermal Decomposition of Dinitrotoluenes. Intermediate Nitroxide Radicals and Reaction Kinetics," submitted to J. Phys. Chem., 1988.

9. Buncel, E., Norris, A. R., and Russell, K. E., "The Interaction of Aromatic Nitro-compounds with Bases," Q. Rev. Chem. Soc., 22 (1968) 123-146.

10. Biswas, M. K., Jarczewski, A., and Leffek, K. T., "Kinetics and Mechanisms of the Reaction of 2,4,6-Trinitrocumene and 2,4,6-Trinitroethylbenzene with 1,1,3,3-tetramethylguanidine in N,N-Dimethylformamide Solvent," Can. J. Chem., 65 (1987) 1007-1011.

11. Andrews, L. J. and Keefer, R. M., Molecular Complexes in Organic Chemistry, Holden-Day, San Francisco, 1964.

12. Anderson, D. R., O'Neill, M. V., and Swanson, J. T., "A Computer Interface to an EPR Spectrometer Using the ASYST Scientific Programming Language," manuscript in preparation for submission to J. Chem. Ed.

USAF-UES MINI-GRANT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

Final Report

Stellar Scintillometer Based Studies of Optical Turbulence

Prepared by:	Frank P. Battles
Academic Rank:	Professor
Department and University:	Basic Sciences Massachusetts Maritime Academy

Date:	December 12, 1987
Contract Number:	F49620-85-C-0013/SB5851-0360

Stellar Scintillometer Based Studies of Optical Turbulence

by

Frank P. Battles

Abstract

Data for C_n^2, the refractive index structure constant, in the approximate altitude range of 2-17 KM was obtained at Pennsylvania State University for the evenings of April 30 - May 6, 1986 using a stellar scintillometer. We here report two major studies which have been done in conjunction with this program. In Chapter 1 we look at correlations of C_n^2 between the different altitude ranges sampled by the scintillometer. For the entire data set significant correlations are obtained. This is not true on a nightly basis. We conclude that these correlations are not instrument caused, but are rather due to the structure of the atmosphere. In Chapter 2 we look at the effects of the differing atmospheric conditions present during the measurement program on θ_o, the isoplanatic angle. We conclude that the presence of jet stream activity strongly effects this parameter.

Acknowledgements

I would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of my research. Special thanks are due to Edmund Murphy of the Optical Physics Division of the Air Force Geophysical Laboratory at Hanscom Field. Both areas of research reported here were suggested by him and his assistance in this research was most helpful. Dr. Laura Kelleher of the Massachusetts Maritime Academy is to be thanked for her valuable input in these endeavours.

Introduction

I received my Ph.D. from Tufts University in theoretical physics in 1969 at which time I joined the faculty at Massachusetts Maritime Academy. In the winter of 1983, while on sabbatical leave, I was employed by Bedford Research on a contract from AFGL. This work involved using a model for the refractive index structure parameter, C_n^2. Using data from the National Weather Service, we looked for seasonal, latitudinal and other variations in this parameter.

During the summer of 1986 I was the recipient of a fellowship in the USAF-UES Summer Faculty Research Program, and was assigned to the Optical Physics Division of AFGL. We looked at methods of calculating θ_o, the isoplanatic angle, from stellar scintillometer derived profiles of C_n^2. This report is a continuation of these studies.

Chapter 1 is entitled *Correlation Coefficients for C_n^2 Based on Stellar Scintillometer Measurements* and has been published as AFGL Technical Memorandum No. 135 with Edmund Murphy as a co-author. Chapter 2 is entitled *The Effects of Meteorological Conditions on Isoplanatic Angle at Pennsylvania State University* and has been published as AFGL Technical Memorandum No. 144 with Laura Kelleher and Edmund Murphy as co-authors. We are in the process of merging these two reports with a third: *Isoplanatic Angle from Stellar Scintillometer Measurements at Pennsylvania State University* (AFGL Technical Memorandum No. 98, 1986, by Battles and Murphy) for the purpose of journal publication.

In the Appendix is a summary of how the grant funding was used.

Chapter 1

Correlation Coefficients for C_n^2 Based on Stellar Scintillometer Measurements

INTRODUCTION

In late April and early May of 1986 Air Force Geophysical Laboratory (AFGL) personnel from the Atmospheric Optics Branch participated in a coordinated field program at Pennsylvania State University (PSU). The purpose of this program was, in part, to compare altitude profiles of C_n^2, the refractive index structure parameter (Hufnagel, 1978) obtained from different measuring devices and to use the acquired data for checking the reliability of certain models for C_n^2. Profiles of this parameter can be obtained using a stellar scintillometer. The one used at PSU by AFGL is a modification (Murphy and Sheldon, 1984) of that originally developed by Ochs et al. (1976). This instrument measures the variance of stellar intensity for a first magnitude or greater star. This variance, by use of spatial filters, can be converted into a spatially and temporally averaged profile of C_n^2 for 7 slant path levels where, by slant path level, we mean distance measured along the line of sight from the instrument to the chosen star. In that the azimuth angle for the star used is known as a function of time, the slant path distances can be converted to obtain the corresponding altitude above ground (Battles and Murphy, 1986).

The C_n^2 values obtained for adjacent levels are not entirely instrument independent because there will be some overlap in the weighting functions used to separate the spatial frequencies. It is usually assumed that those readings from levels 1,4 and 7 are instrument independent.

It has been brought to the attention of AFGL by their co-investigators in this measurement program from Rome Air Development Center(RADC) that correlation coefficients for C_n^2 values derived from AFGL data (192 profiles) between different levels seem surprisingly large. They comment that this phenomena is true for other scintillometer data sets from other measurement programs and suggest that this may be instrument related.

RADC also used a stellar scintillometer at PSU but their data (140 profiles) does not lead to such a high level of correlation.

In the absence of any other considerations, the correlation coefficients calculated by RADC for the AFGL scintillometer data could indicate that, for all levels, C_n^2 values are instrument dependent contradicting the above stated assumption. It is the purpose of this report to further consider these observations and to suggest other possible reasons for the high degree of correlation observed.

CALCULATIONS

Data is available for the evenings of April 30 and May 2-6. The number of profiles for each evening are 12, 39, 56, 26, 16, and 43 respectively, a total of 192.

We began by looking at the difference between slant path levels and altitude above ground for the AFGL scintillometer profiles. During the PSU program, the instrument was pointed at the same star (Arctaurus) at a near zenith position. The actual altitude for each measurement was calculated. Then the mean and standard deviation for all measurements were calculated. (See Table 1). In that the standard deviation is of the order of 5% of the mean for each slant path level, we shall assume that the C_n^2 values from a given level correspond to the mean altitudes presented in Table 1.

We next calculated (linear) correlation coefficients (see for example, Brockett and Levine, 1984) between each level for both C_n^2 and log C_n^2 for each evening and for the entire program. These results are presented in Tables 2-8. Each table presents, in matrix form, 2 sets of correlation coefficients: $r_n(i,j)$ and $r_L(i,j)$ for the correlations of the numerical and logarithmic values of C_n^2 between level i and j. Tables 2-7 are for each evening and Table 8 contains results for the entire measurement program.

Our results for the 192 AFGL profiles differ slightly for C_n^2 (but not log C_n^2) from those obtained by RADC for the same data set. Although differences are small (as an extreme we calculate $r_n(4,5)$ to be 0.87 whereas RADC calculations yield 0.77), this is puzzling. For this reason a copy of the program used is included as Appendix 1. A possible explanation may be that in the statistical calculations products of C_n^2 may have been inadvertently set equal to zero by whatever calculating device was used by RADC. Using GW-BASIC software on a Zenith 158 will cause this to happen for numbers less than 10^{-38}. To check that this was not happening in our calculations we recalculated after pre-multipling C_n^2 by 10^6 and found no difference.

When we look at the results for all 192 profiles we see for both C_n^2 and log C_n^2 a significant positve correlation between all levels. Correlation coefficients decrease as the altitude separation between levels increases. Level 1(at 2km) and level 2(at 3.1 km) are the closest and show the largest correlation. Levels 1 and 7(at 16.6km) show the least.

For the entire data set a much different picture emerges than that of any individual evening except for low lying adjacent layers. Consider, for example, $r_n(1,7)$ which for successive evenings takes on the values 0.44 (12 profiles), 0.08(39 profiles), 0.27(56 profiles),0.18(26 profiles), -0.12(16 profiles) and 0.25(43 profiles), yielding a weighted mean of 0.19. For all 192 profiles $r_n(1,7) = 0.37$, i.e. about a factor of two larger than nightly measurements would suggest. Similarly for $r_n(6,7)$ we have on a nightly basis -0.01, 0.24, 0.60, 0.47, 0.03, and 0.04 giving a weighted mean of 0.30 whereas for the entire sample we have $r_n(6,7) = 0.60$.

Statistically significant negative correlations have been obtained on some evenings. For example, for May 2 based on 39 profiles $r_L(1,4) = -0.31$ which yields at t-statistic (Brockett and Levine, 1984) of -1.98 which is

significant at the 0.05 level. For May 5 based on 26 profiles we find that $r_L(1,4) = 0.31$ which yields a t-statistic (Brockett and Levine, 1984) of -1.98 which is significant at the 0.05 level. For May 5 based on 26 profiles we find that $r_L(3,7) = -0.49$ which yields a t-statistic of -2.75 which is significant at the 0.01 level.

Scatter plots have been produced to illustrate the above discussion. In Figures 1 and 2 are shown graphs of C_n^2 and log C_n^2 for levels 1 versus 7 for the entire data set. We also show the least squares line. Similarly, Figures 3 and 4 are for level 1 versus 2. In Figure 5 we display a typical scatter diagram for a negative correlation coefficient and in Figure 6 such a plot for a near zero correlation.

We also have calculated for C_n^2 and for log C_n^2 correlation coefficients with altitude for the entire data set obtaining values of -0.31 for C_n^2 and -0.50 for log C_n^2.

For each evening and for the entire sample we have calculated the slope and intercept of the least squares line for C_n^2 and log C_n^2. These results are not presented here but will be made available upon request.

DISCUSSION

Various models for C_n^2 have been proposed. In most of these altitude is directly or indirectly the dominant parameter. Any model of the form $C_n^2(z) = C_n^2(z_o)xf(z)$ where f(z) is any function of altitude and $C_n^2(z_o)$ is some reference value would predict a correlation of unity. Some such models include that of Kaimal as modified by Walters and Kunkel (1981), that of Battles, Murphy and Noonan(1987) and that of Kukharets and Tsang(1980), each of which would yield a 100% correlation of C_n^2 from level to level. We are not saying that these models do apply to the altitude range covered by this data but are suggesting that the high degree of correlation observed may be due to the fact that, although very different

physical mechanisms are producing turbulence at the different levels sampled, when sufficient data is used we expect that on the whole C_n^2 will be a function of altitude. This would explain the difference between nightly results versus that observed over the entire measurement program. From night to night differing atmospheric conditions exist. When these are averaged over all six nights C_n^2 becomes correlated due to the fact that it is overall a decreasing function of altitude.

The correlation coefficients calculated for C_n^2 and log C_n^2 with altitude are a further indication that altitude, not instrumentation, is affecting the values of $r_n(i,j)$ and $r_L(i.j)$. Should a model of the above suggested form hold, we would expect that a least squares line for log C_n^2 from level to level would have slope 1. Our calculations for adjacent levels yield slopes of order 0.8.

If the observed correlations were primarily instrument caused, we would expect a similar pattern on any given evening similar to that exhibited overall: statistically significant positive correlation which decreases with level separation. Weighting function overlap can only result in positive correlation, yet significant negative correlation has been observed on several occasions.

CONCLUSIONS

Based on the above stated reasons we do not feel that one can conclude that stellar scintillometer obtained profiles for C_n^2 are instrument dependent for all levels.

FURTHER EFFORTS

We are presently investigating the effects of the synoptic conditions (frontal type present, location relative to the jet stream etc.) which existed on each evening. This may lead to an understanding of the differing correlation patterns observed from evening to evening. These studies might explain the cause of the negative correlations noted above.

Similar night by night versus overall correlation studies for other stellar scintillometer profiles of C_n^2 should be undertaken to see if results are similar to those presented here. In particular it would be most informative to know how RADC scintillometer results obtained at PSU for $r_n(i,j)$ and $r_L(i,j)$ compare to those of AFGL on an evening to evening basis and to see if, over the same period, the same degree of correlation is observed.

ACKNOWLEDGEMENTS

The authors wish to express their thanks to Dr. Laura Kelleher of Massachusetts Maritime Academy (MMA), for much assistance regarding the statistical calculations and interpretation of such. Dr. Kenneth Gucwa (MMA) is to be thanked for his assistance in some of the computational aspects of this report.

REFERENCES

Battles, F. P., E. A. Murphy and J. P. Noonan, "The Contribution of Atmospheric Density to the Drop-off Rate of C_n^2", submitted for Publication, Physica Scripta.

Brockett P. and A. Levine, Statistics and Probability and Their Applications, Chapter 9, Saunders College Publishing, 1984.

Hufnagel, R. E., "Propagation Through Atmospheric Turbulence", in The Infrared Handbook, U. S. Government Printing Office, Washington, D.C., Chapter 6, 1978.

Kukharets, V. P. and L. R. Tsang, "Structure Parameter of the Refractive Index in the Atmospheric Boundary Layer", IZV: Atmospheric and Oceanic Physics, Vol. 16, pp. 397-405, 1980.

Murphy, E. A., R. E. Good and S. Sheldon, "Stellar Scintillometer Measurements of C_n^2 at Kirtland AFB, New Mexico", AFGL Technical Memorandum, No. 98, 1984.

Murphy, E. A. and F. P. Battles, "Isoplanatic Angle from Stellar Scintillometer Measurements at Pennsylvania State University", AFGL Technical Memorandum, No. 129, 1986.

Ochs, G. R., Ting-i Wang, R. S. Lawrence, and S. F. Clifford, "Refractive Turbulence Profiles Measured by One-Dimensional Spatial Filtering", Applied Optics, Vol. 15 pp. 2504-2510, 1976.

Walters D. L. and K. E. Kunkel, "Atmospheric Transfer Function for Desert and Mountain Locations: The Atmospheric Effects on r_o", Journal of the Optical Society of America, Vol. 71, pp. 397-405(1981).

Cn2 Correlations

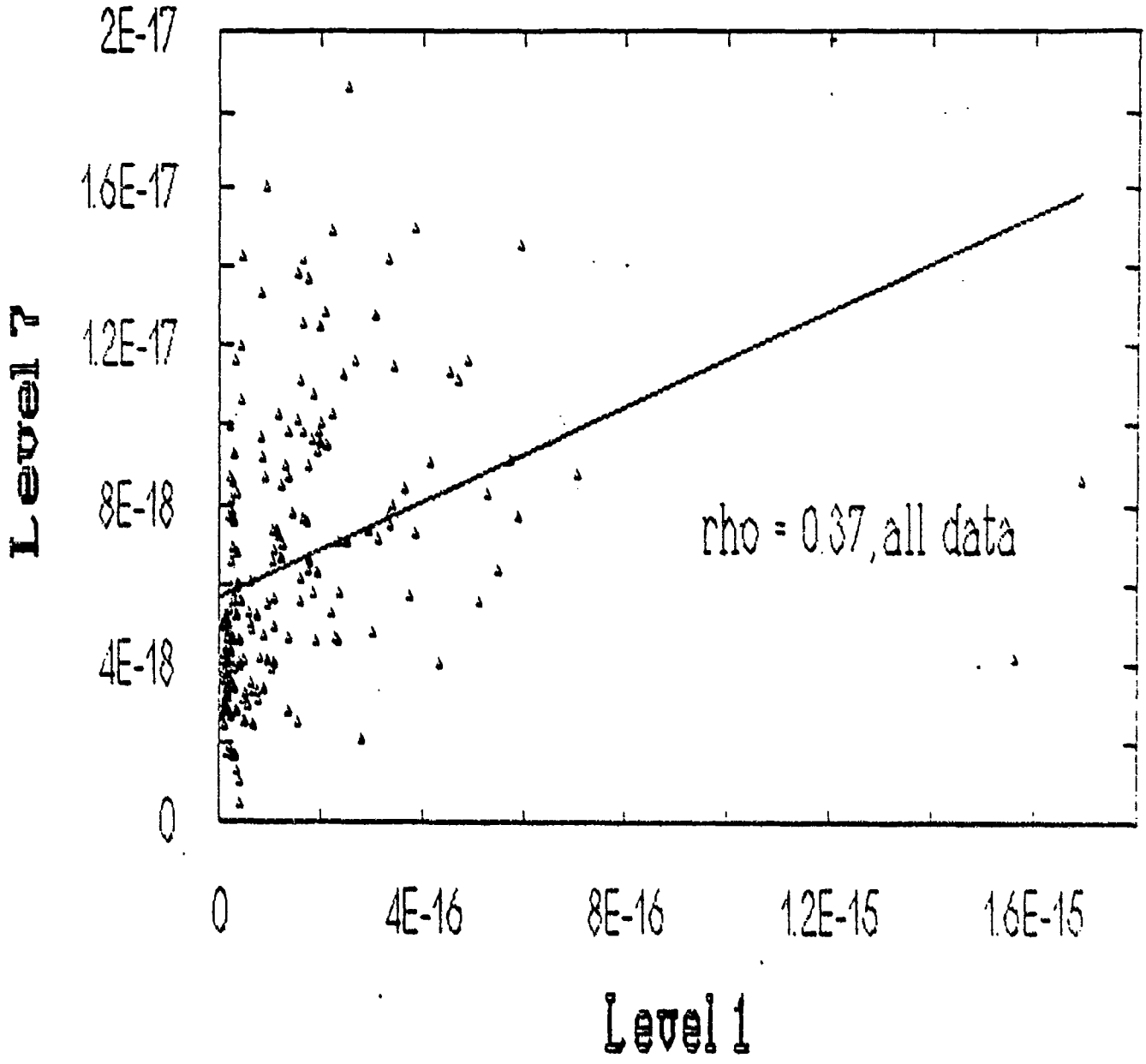


Figure 1

Log Ch2 Correlations

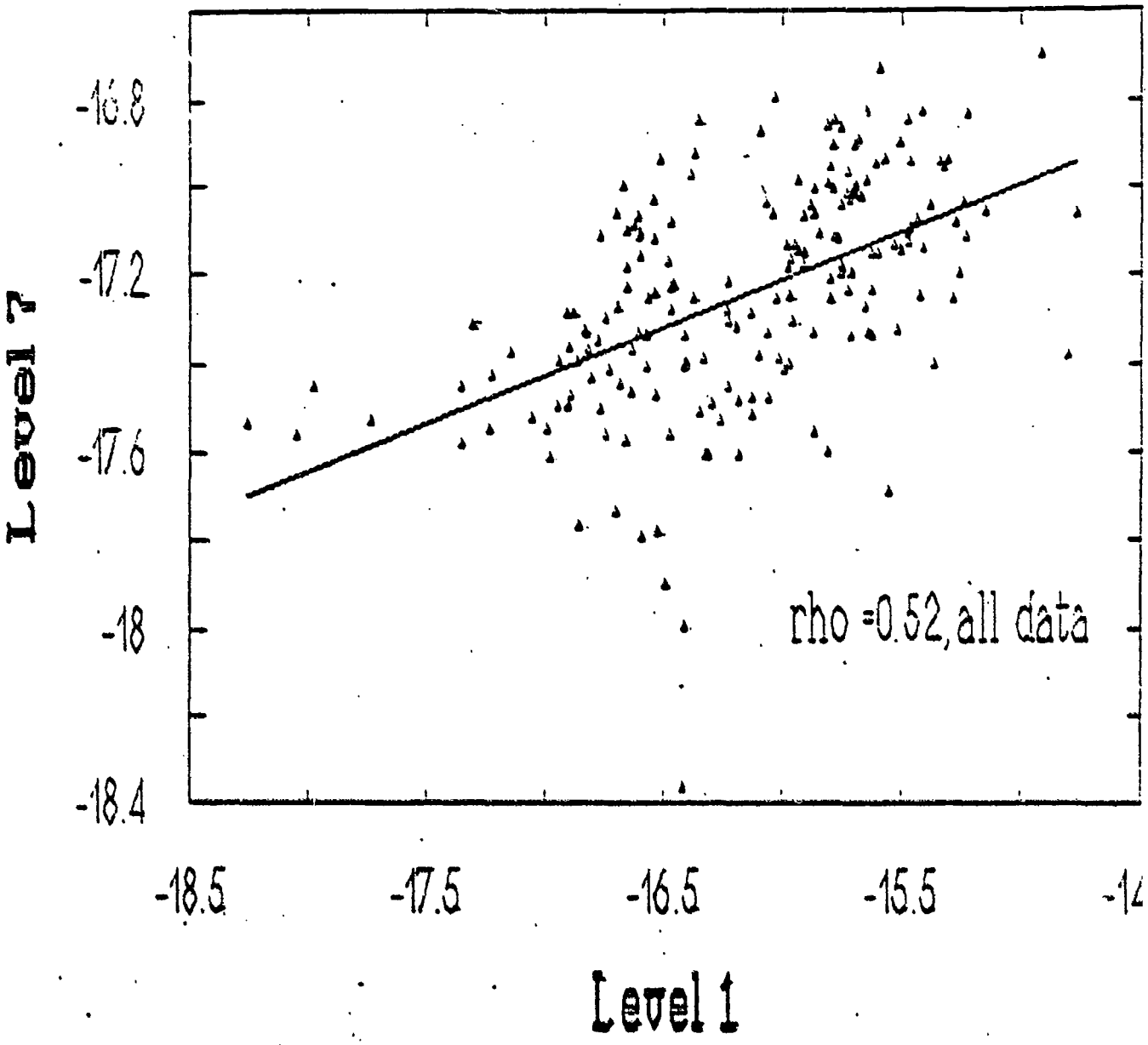


Figure 2

Cn2 Correlations

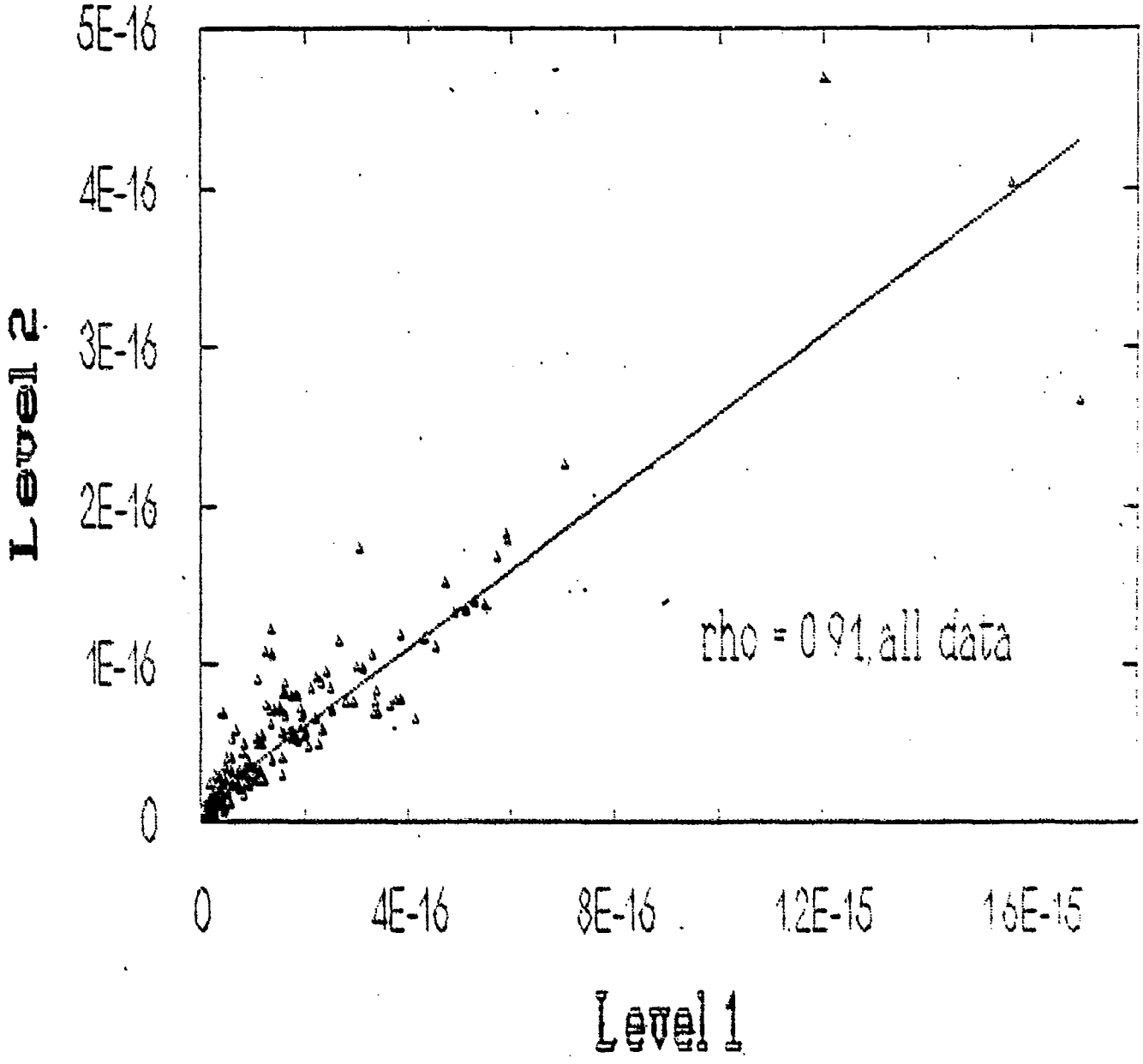


Figure 3

Log Ch2 Correlations

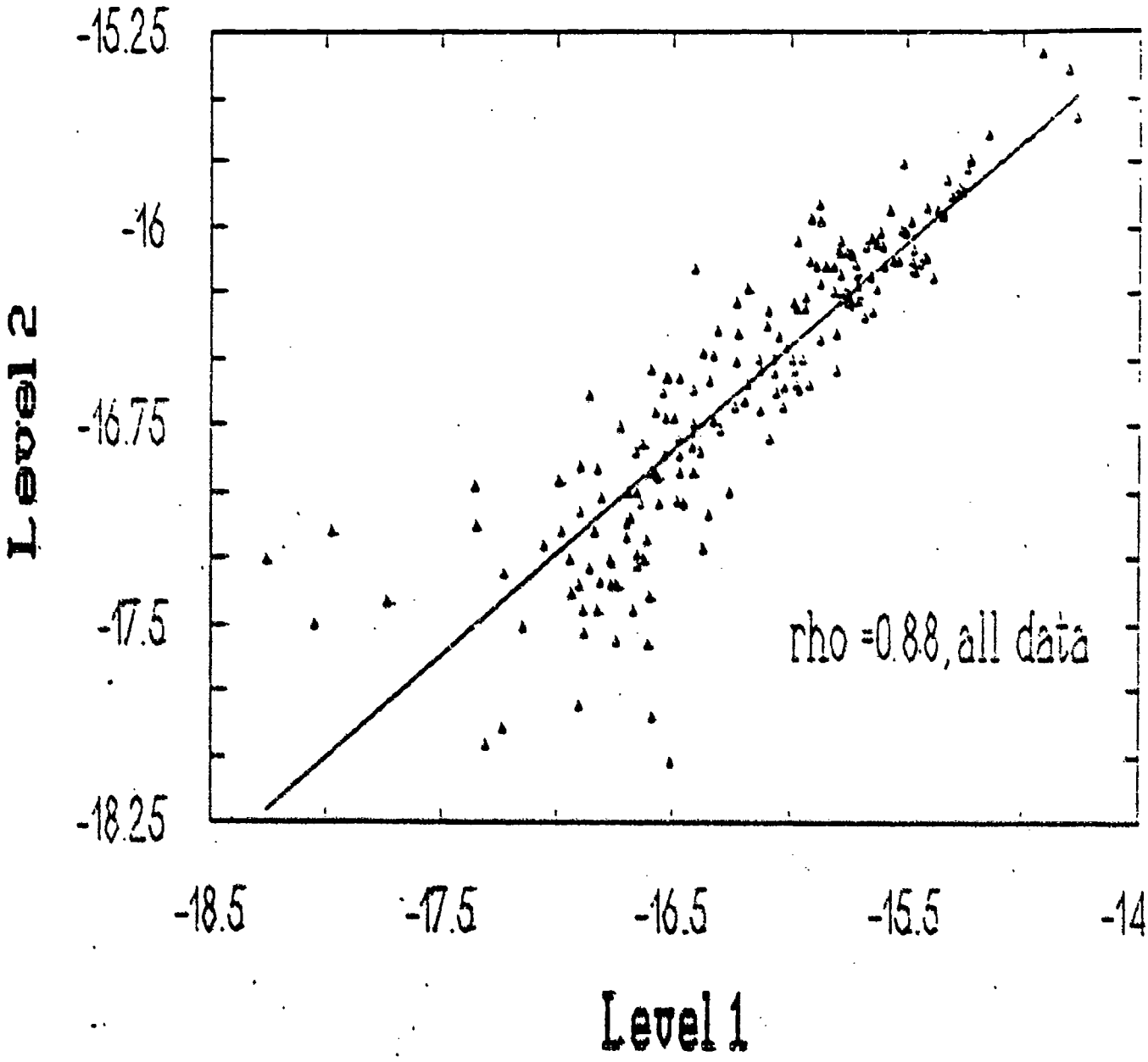


Figure 4

Log Ch2 Correlations

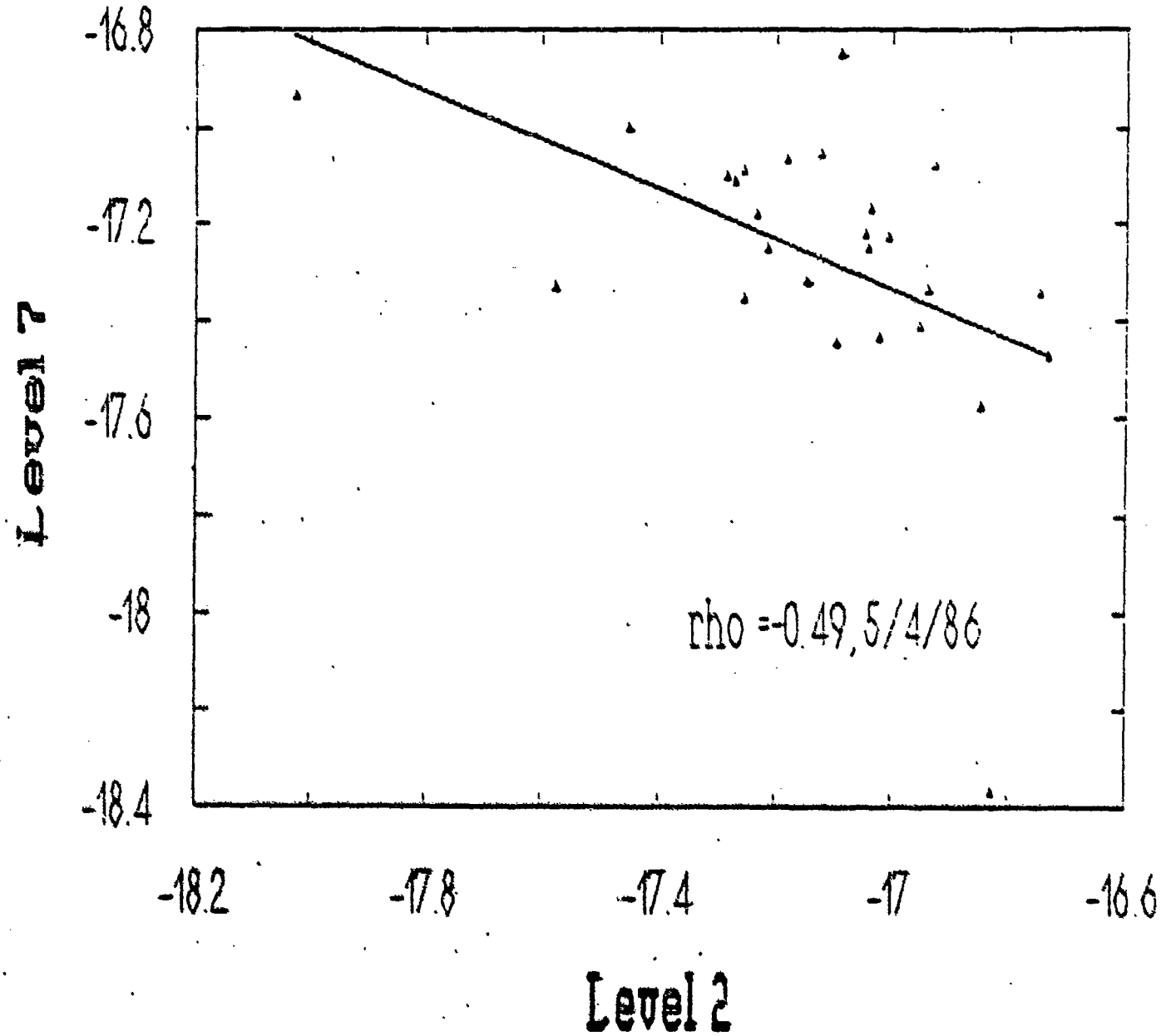


Figure 5

Ch2 Correlations

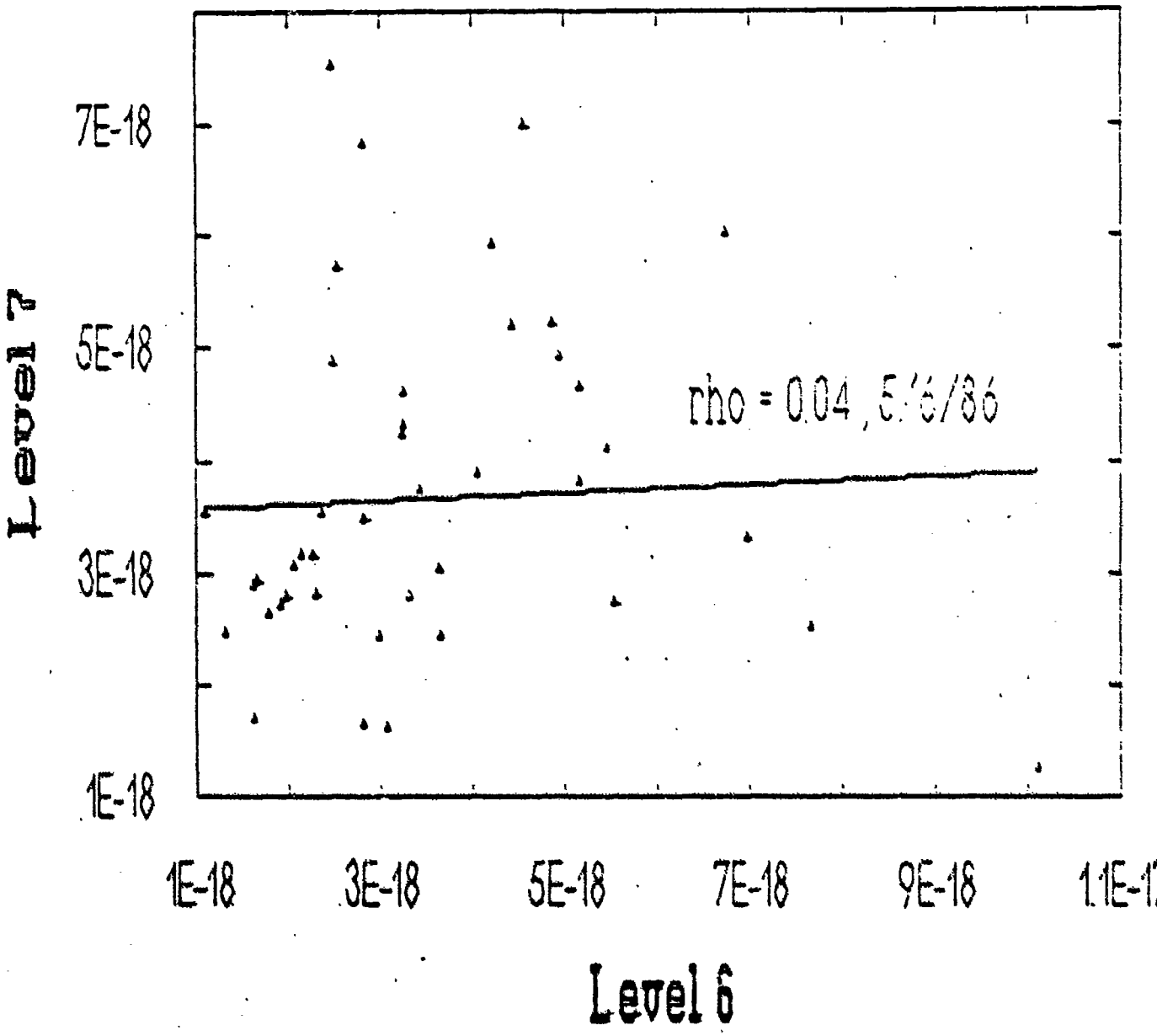


Figure 6

BASED ON 192 SAMPLES

THE AVERAGE ALTITUDE AT LEVEL 1 IS 1.97KM WITH STANDARD DEVIATION 0.11KM

THE AVERAGE ALTITUDE AT LEVEL 2 IS 3.05KM WITH STANDARD DEVIATION 0.17KM

THE AVERAGE ALTITUDE AT LEVEL 3 IS 4.66KM WITH STANDARD DEVIATION 0.26KM

THE AVERAGE ALTITUDE AT LEVEL 4 IS 6.54KM WITH STANDARD DEVIATION 0.36KM

THE AVERAGE ALTITUDE AT LEVEL 5 IS 8.42KM WITH STANDARD DEVIATION 0.46KM

THE AVERAGE ALTITUDE AT LEVEL 6 IS 12.54KM WITH STANDARD DEVIATION 0.69KM

THE AVERAGE ALTITUDE AT LEVEL 7 IS 16.57KM WITH STANDARD DEVIATION 0.91KM

Table 1

Correlation coefficients between layers

AFGL data (Penn State) for numerical values

Data for April 30, 1986. 12 profiles

	1	2	3	4	5	6	7
1		0.96	0.92	0.51	0.70	-0.04	0.44
2			0.97	0.36	0.61	0.05	0.38
3				0.43	0.62	0.26	0.39
4					0.86	0.11	0.67
5						-0.04	0.70
6							-0.01
7							

Correlation coefficients between layers

AFGL data (Penn State) for logarithmic values

Data for April 30, 1986. 12 profiles

	1	2	3	4	5	6	7
1		0.96	0.95	0.72	0.70	0.18	0.67
2			0.99	0.57	0.63	0.15	0.54
3				0.61	0.66	0.22	0.50
4					0.86	0.09	0.58
5						-0.04	0.62
6							0.02
7							

Table 2

Correlation coefficients between layers

AFGL data (Penn State) for numerical values

Data for MAY 2, 1986. 39 profiles

	1	2	3	4	5	6	7
1		0.86	0.43	-0.26	-0.25	0.17	0.08
2			0.71	-0.17	-0.18	0.26	0.12
3				0.43	0.28	0.34	0.06
4					0.88	0.05	0.14
5						-0.06	0.24
6							0.24
7							

Correlation coefficients between layers

AFGL data (Penn State) for logarithmic values

Data for MAY 2, 1986. 39 profiles

	1	2	3	4	5	6	7
1		0.81	0.44	-0.31	-0.28	0.04	0.09
2			0.72	-0.31	-0.31	0.15	0.10
3				0.22	0.08	0.29	0.04
4					0.90	-0.12	0.05
5						-0.15	0.15
6							0.23
7							

Table 3

Correlation coefficients between layers

AFGL data (Penn State) for numerical values

Data for MAY 3,1986. 56 profiles

	1	2	3	4	5	6	7
1		0.92	0.85	0.39	0.30	0.24	0.27
2			0.93	0.42	0.25	0.46	0.38
3				0.65	0.42	0.31	0.34
4					0.87	0.10	0.36
5						0.12	0.34
6							0.60
7							

Correlation coefficients between layers

AFGL data (Penn State) for logarithmic values

Data for MAY 3,1986. 56 profiles

	1	2	3	4	5	6	7
1		0.89	0.70	0.16	0.20	0.12	0.28
2			0.90	0.35	0.35	0.24	0.33
3				0.62	0.53	0.24	0.49
4					0.83	0.19	0.43
5						0.22	0.34
6							0.47
7							

Table 4

Correlation coefficients between layers

AFGL data (Penn State) for numerical values

Data for MAY 4, 1986. 26 profiles

	1	2	3	4	5	6	7
1		0.33	-0.03	-0.07	0.20	0.29	0.18
2			0.73	0.34	0.06	-0.14	-0.55
3				0.65	0.11	-0.24	-0.54
4					0.41	-0.03	-0.15
5						0.35	0.17
6							0.47
7							

Correlation coefficients between layers

AFGL data (Penn State) for logarithmic values

Data for MAY 4, 1986. 26 profiles

	1	2	3	4	5	6	7
1		0.19	-0.06	0.01	0.21	0.17	-0.05
2			0.57	0.39	-0.04	-0.18	-0.49
3				0.81	0.21	-0.19	-0.44
4					0.45	0.05	-0.13
5						0.23	-0.03
6							0.17
7							

Correlation coefficients between layers

AFGL data (Penn State) for numerical values

Data for MAY 5, 1986. 16 profiles

	1	2	3	4	5	6	7
1		0.53	0.40	0.14	-0.23	0.25	-0.12
2			0.93	0.30	-0.13	0.20	-0.13
3				0.40	-0.03	0.15	-0.08
4					0.53	0.04	0.10
5						-0.12	0.45
6							0.03
7							

Correlation coefficients between layers

AFGL data (Penn State) for logarithmic values

Data for MAY 5, 1986. 16 profiles

	1	2	3	4	5	6	7
1		0.54	0.39	0.26	-0.37	0.24	-0.17
2			0.89	0.13	-0.16	0.17	-0.11
3				0.10	-0.14	0.12	0.03
4					0.06	-0.01	0.06
5						-0.11	0.45
6							0.03
7							

Table 6

AFGL data (Penn State) for numerical values

Data for MAY 6, 1986. 43 profiles

	1	2	3	4	5	6	7
1		0.88	0.85	0.80	0.83	0.50	0.25
2			0.99	0.86	0.64	0.28	0.27
3				0.89	0.62	0.26	0.28
4					0.71	0.40	0.27
5						0.63	0.12
6							0.04
7							

Correlation coefficients between layers

AFGL data (Penn State) for logarithmic values

Data for MAY 6, 1986. 43 profiles

	1	2	3	4	5	6	7
1		0.77	0.72	0.52	0.66	0.66	0.17
2			0.88	0.73	0.62	0.43	0.21
3				0.88	0.68	0.38	0.20
4					0.69	0.36	0.13
5						0.57	0.04
6							0.08
7							

Table 7

Correlation coefficients between layers

AFGL data (Penn State) for numerical values

Data for April30-May 6, 1986. 192 profiles

	1	2	3	4	5	6	7
1		0.91	0.82	0.44	0.42	0.40	0.37
2			0.93	0.52	0.45	0.52	0.43
3				0.72	0.57	0.44	0.40
4					0.87	0.34	0.46
5						0.42	0.56
6							0.60
7							

Correlation coefficients between layers

AFGL data (Penn State) for logarithmic values

Data for April30-May 6, 1986. 192 profiles

	1	2	3	4	5	6	7
1		0.98	0.76	0.54	0.61	0.62	0.52
2			0.92	0.66	0.64	0.53	0.43
3				0.77	0.68	0.49	0.41
4					[illegible]	0.42	0.38
5						0.54	0.49
6							0.55
7							

Table 8

```
10 LPRINT"Correlation coefficients between layers"
20 LPRINT
30 LPRINT"AFGL data (Penn State) for numerical values"
40 LPRINT
50 DIM C(7)
60 DIM SX(7,7),SY(7,7),SXSQ(7,7),SYSQ(7,7),SXY(7,7)
70 FOR J=1 TO 7
80 FOR K=1 TO 7
90 SX(J,K)=0:SY(J,K)=0:SXSQ(J,K)=0:SYSQ(J,K)=0:SXY(J,K)=0
100 NEXT K
110 NEXT J
120 COUNT=0
130 DIM RHO(7,7)
140 DIM NUM(7,7),DEN(7,7)
150 REM SCINT2.DAT HAS DATA FOR ALL 192 PROFILES
160 OPEN "I",#1,"A:scint2.DAT"
170 INPUT#1,D,M,HR,MN,SEC,C(1),C(2),C(3),C(4),C(5),C(6),C(7)
180 REM FOR J=1 TO 7:C(J)=LOG(C(J)):NEXT J
190 FOR J=1 TO 6
200 FOR K=(J+1) TO 7
210 SX(J,K)=SX(J,K)+C(J)
220 SY(J,K)=SY(J,K)+C(K)
230 SXSQ(J,K)=SXSQ(J,K)+C(J)^2
240 SYSQ(J,K)=SYSQ(J,K)+C(K)^2
250 SXY(J,K)=SXY(J,K)+C(J)*C(K)
260 NEXT K
270 NEXT J
280 COUNT=COUNT+1
290 IF NOT EOF(1) THEN GOTO 170
300 LPRINT"Data for April30-May 6,1986.";COUNT;"profiles"
310 LPRINT
320 CLOSE
330 LPRINT "       1       2       3       4       5       6       7"
340 LPRINT:LPRINT
350 FOR J=1 TO 6
360 LPRINT J;
370 FOR L=1 TO 8*J
380 LPRINT" ";
390 NEXT L
400 FOR K=(J+1) TO 7
410 NUM(J,K)=COUNT*SXY(J,K)-SX(J,K)*SY(J,K)
420 DEN(J,K)=((COUNT*SXSQ(J,K)-SX(J,K)^2))^.5*((COUNT*SYSQ(J,K)-SY(J,K)^2))^.5
430 RHO(J,K)=NUM(J,K)/DEN(J,K)
440 LPRINT USING" #.##  ";RHO(J,K);
450 NEXT K
460 LPRINT:LPRINT:LPRINT
470 NEXT J
480 LPRINT" 7"
490 END
```

Chapter 2

The Effects of Meteorological Conditions on Isoplanatic Angle at Pennsylvania State University

INTRODUCTION

As an optical beam propagates through the atmosphere, fluctuations in the index of refraction will induce distortions which degrade the optical quality of the beam. One measure of these distortions is the isoplanatic angle. We have examined the effects of various meteorological conditions, especially jet stream activity, on this parameter.

BACKGROUND

In late April and early May of 1986 Air Force Geophysical Laboratory (AFGL) personnel from the Atmospheric Optics Branch participated in a coordinated field program at Pennsylvania State University(PSU). The purpose of this program was, in part, to compare altitude profiles of C_n^2, the refractive index structure parameter (Hufnagel,1978) obtained from different measuring devices and to use the acquired data for checking the reliability of certain models for C_n^2. Profiles of this parameter can be obtained using a stellar scintillometer. The one used at PSU by AFGL is a modification (Murphy et. al.,1984) of that originally developed by Ochs et al.(1976.) This instrument measures the variance of stellar intensity for a first magnitude or greater star. This variance, by use of spatial filters, can be converted into a spatially and temporally averaged profile of C_n^2 for seven slant path levels where, by slant path level, we mean distance measured along the line of sight from the instrument to the chosen star. In that the azimuth angle for the star used is known as a function of time, the slant path distances can be converted to obtain the corresponding altitude above ground (Murphy and Battles, 1986.) A total of 192 profiles were obtained : 12 for April 30, 39 for May 2, 56 for May 3, 26 for May 4, 16 for May 5 and 43 for May 6. A typical profile is shown in Figure 1.

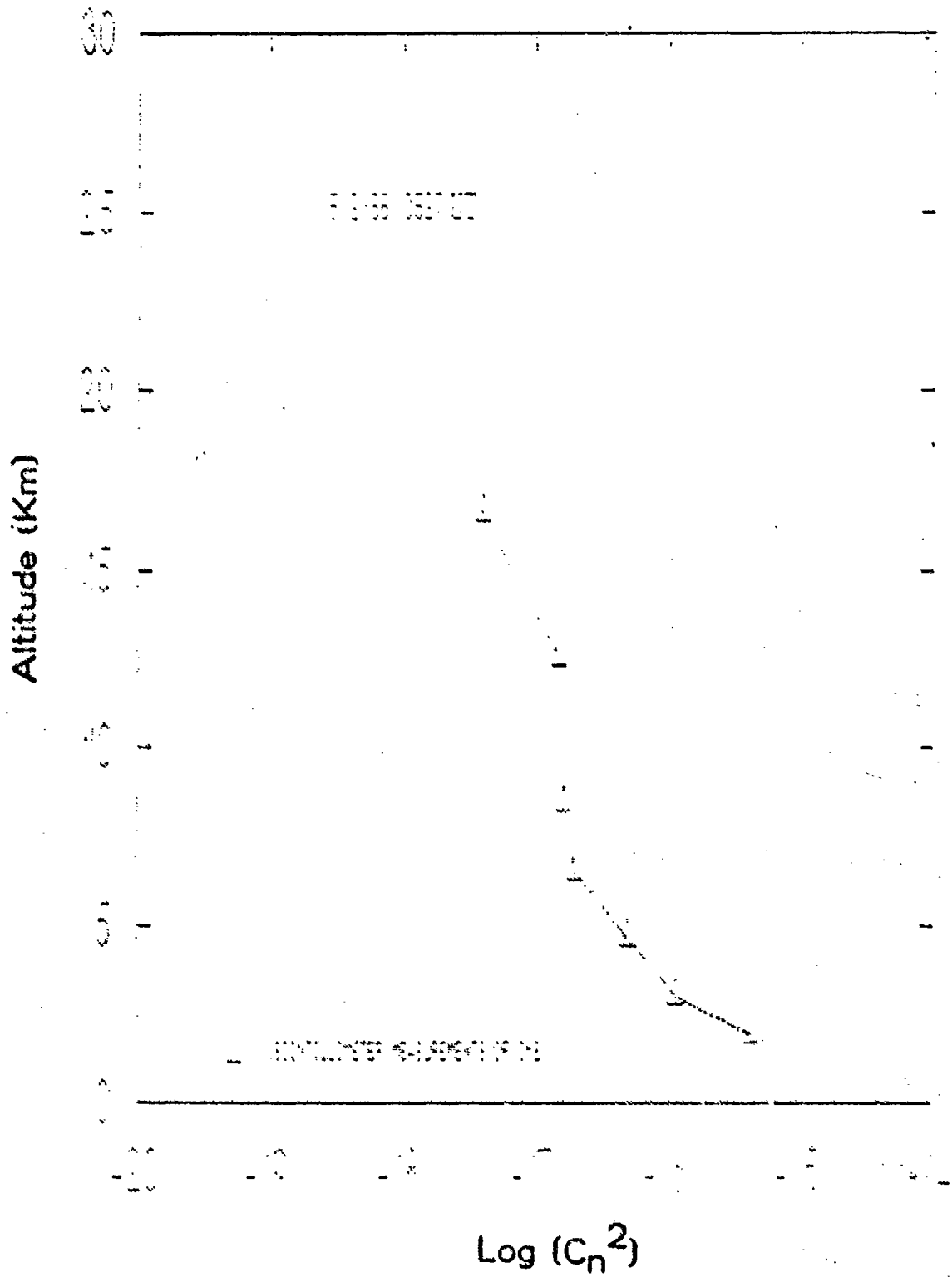


FIGURE 1

Isoplanatic Angle

One important measure of turbulence induced distortions on an optical beam is the θ_0, the isoplanatic angle (Born and Wolf, 1983) which is the maximum angle over which an extended object can be viewed, in the absence of adaptive optics, through turbulence. For the case of vertical viewing down through the atmosphere Loos and Hogge (1979) have shown that, for θ_0 in microradians,

$$\theta_0 = 5.8 \times 10^4 \lambda^{1.2} \left[\int_0^\infty z^{\frac{5}{3}} C_n^2(z)\, dz \right]^{-\frac{3}{5}}$$

where λ = wavelength(m)

C_n^2 = Refractive index structure parameter($m^{-2/3}$)

z = altitude above ground(m).

The effects of the $z^{5/3}$ weighting of C_n^2 in the above integral is such that contributions from that portion of the atmosphere covered by the scintillometer is dominant. This allows one to use models for the regions above (Battles et. al.,1987) and below (Walters and Kunkel, 1981 and Kukharkets and Tsang, 1990) the scintillometer range for C_n^2 and obtain a reasonable estimate for θ_0 . Good agreement with isoplanometer measurements of θ_0 has been obtained as well as with values calculated from thermosonde profiles of C_n^2 (Murphy and Battles, 1987.) We further note that this weighted integration of C_n^2 should amplify the effects of changing atmospheric conditions in the scintillometer range.

It is the purpose of this report to investigate the effects of the meteorological conditions present during the PSU program on θ_0, especially the effects of jet stream passages. We report in our calculations a wavelength of 500 nm and will report values of log θ_0 for θ_0 in microradians.

The Meteorological Parameters

A set of basic meteorological parameters was supplied to us by the Meteorology Department of PSU. Values for these parameters were obtained from local weather conditions and twice daily weather maps at 200 and 300 millibars. Interpolation and extrapolation were used to provide estimates of these parameters on an hourly basis from April 28 to May 6, 1986. In all, data was given for some fourteen parameters. The data was coded in the following fashion:

Column 1. Circulation type(surface)
1 = low(cyclonic), 2 = high(anticyclonic), 3 = neither
Column 2. Location relative to the circulation center(surface)
1 = NW, 2 = SW, 3 = SE, 4 = NE, 5 = circulation center
Column 3. Frontal type (surface)
1 = warm, 2 = cold, 3 = occluded, 4 = no front present
5 = trough
Column 4. Location relative to the front(surface)
1 = warm side, 2 = cold(dry) side, 3 = within frontal zone, 4 = not near front(not within 300km of front, 5 = ahead, 6 = behind(5 or 6 if occluded or trough)
Column 5. Wave category (upper level)
1 = northerly wind maximum, 2 = trough, 3 = southerly wind maximum, 4 = ridge, 5 = zonal front, 6 = split flow center(very weak height gradient), 7 = cutoff low within 300km(two or more closed contours at 200 or 300 mb), 8 = light and variable flow(very weak height gradient)
Column 6. Position relative to jet streak
1 = left front, 2 = right front, 3 = left rear, 4 = right rear, 5 = no streak present
Column 7. Frontal type(upper level)
1 = cold, 2 = warm, 3 = occluded, 4 = no front present
Column 8. Location relative to the front(upper level)
1 = ahead, 2 = behind, 3 = within frontal zone, 4 = not near front
Column 9. Position relative to the jet axis
1 = left(0-150 km), 2 = right(0-150 km), 3=left(150-300km) 4=right(150-300km) 5=greater than 300km 6=jet axis 9=jets of equal strength right and left of station, neither jet dominates.
Column 10. Cloud type
1 = clear, 2 = shallow convection, 3 = deep convection 4 = low stratefors, 5 = middle, 6 = high, 7 = layered cloud
Column 11. Position relative to cloud area
1 = NW(0-150km), 2 = NE(0-150km), 3 = SE(0-150km), 4 = SW(0-150km), 5 = NW(150-300km), 6 = NE(150-300km), 7 = SE(150-300km), 8 = SW(150-300km), 9 = not applicable(either no cloud areas within 300km or station under clouds of type determined from entry 10), 11 = no cloud area within 300km(if clouds reported at station), 22 = cloud areas in two or more quadrants within 300km

Column 12. Mesoscale influence
1 = cold air damming, 2 = mesoscale convective complex, 3 = squall line, 4 = tropical disturbance, 5 = none detected, 6 = mesohigh, 7 = mesolow(6 or 7 indicate the presence of a thunderstorm complex, but possibly not a squall line or mesoscale convective complex; when satellite data unavailable for proof, 6 or 7 is all that can be entered
Column 13. Hourly wind direction at 200mb, in tens of degrees
Column 14. Hourly wind direction at 300mb, in tens of degrees

Note : throughout a value of nine indicates not applicable and zero indicates not available.

Calculations

We first established a data file consisting of the 192 values of θ_e obtained from the scintillometer profiles, the date and time the profile was obtained and the values of the fourteen parameters for this time period. In that over this time period no jet streak was present and no mesoscale influence was detected, parameters =6 and =12 were eliminated from further study. For the remaining parameters, we next formed bins of those values of θ_e corresponding to the same value of a given parameter. Within each bin the mean and standard deviation of log θ_e was calculated. These results are presented in Table 1 which is on the next page. Column 1 is the number of the parameter, column 2 the value of the parameter, column 3 the number of profiles in the associated bin, column 4 gives the mean value of log θ_e, and column 5 is the standard deviation of log θ_e.

The log mean was chosen in that it is usually assumed that C_n^2 and θ_e are log normally distributed (Loos and Hogge,1979.) We have for the overall distribution of θ_e calculated the mean, median, standard deviation, skewness, and kirtosis obtaining values of 7.6, 7.0, 2.5, 0.6 and -0.5 respectively. The discrepancy of the mean and median and the magnitude of the skewness would indicate a non-normal distribution for θ_e. For log θ_e the above statistical parameters take on the values 0.86, 0.85, 0.14, 0.04 and -0.36. These results are consistent with the assumption of a log normal distribution (Scheffe, 1959).

PARAMETER	VALUE	COUNT	LOG θ_0	S.D.
1	2	133	0.80	0.11
1	3	59	1.00	0.09
2	3	95	0.77	0.10
2	4	38	0.87	0.10
2	9	59	1.00	0.09
3	1	17	1.04	0.06
3	2	42	0.99	0.10
3	4	94	0.84	0.10
3	5	39	0.70	0.06
4	1	59	1.00	0.09
4	4	94	0.84	0.10
4	5	39	0.70	0.06
5	1	97	0.95	0.11
5	2	95	0.77	0.10
7	1	56	0.81	0.10
7	2	65	0.78	0.12
7	4	71	0.97	0.12
[illegible]	3	121	0.79	0.11
[illegible]	4	71	0.97	0.12
9	1	57	0.85	0.11
9	2	39	0.70	0.06
9	5	59	1.00	0.09
9	6	37	0.81	0.09
10	1	182	0.85	0.14
10	6	10	1.05	0.05
11	0	29	0.95	0.14
11	3	10	0.72	0.04
11	7	29	0.69	0.06
11	9	124	0.89	0.12
13	27	95	0.77	0.10
13	31	51	0.94	0.11
13	32	29	0.92	0.13
13	33	17	1.04	0.06
14	27	95	0.77	0.10
14	30	42	0.99	0.10
14	32	33	0.92	0.10
14	33	22	0.92	0.15

TABLE I

(for λ = 500 nm and θ_0 in microradians)

Discussion·

A t-statistic (Brockett and Levine, 1984) between differing values of the same parameter was calculated. Many statistically significant differences were found across bins for the same parameter. This does not necessarily indicate causality. We note that many of the binnings of the data by parameter number and value are not independent. For example, parameter 9, value 5 (jet stream absent) is the same bin as parameter 1, value 3 (surface circulation type: none), parameter 2 value 9 (location relative to circulation n/a) and parameter 4, value 1(location relative to surface front: warm side.)

Several jet stream passages were noted during the measurement program. In that one of the purposes of this program was to investigate the effects of the jet stream on C_n^2 and hence log Θ, we have focused our attention on parameter #9 and the times for these passages. Comparing weighted means calculated from the data of Table I, we see that log Θ is significantly larger when the jet stream is absent (value = 5) with a mean of 1.0 than when the jet stream is either directly overhead (value = 6) or within 150km to the left (value = 1) or the right (value = 2). These later three values yield a weighted mean of 0.79.

From Table 1, parameter #9 we see that it would appear as if there were a much more pronounced effect in having the jet stream to the right of the measurement station than to the left or overhead. A more detailed analysis of the 200 mbar and 300 mbar weather maps may explain this. On April 30, May 2 and May 3 the jet stream was either overhead, or in motion toward the station. The evening of May 4 corresponds to a transition period with the stream moving away from the station. The evenings of May 5 and May 6 correspond to an absence of the jet stream. This suggests a different binning of the data: i.e.

1. Jet stream active (April 30, May 2, May 3)

2. Jet stream absent (May 5, May 6)

3. Transition period (May 4)

The results of this grouping are shown in Table 2 below.

	Sample size	Mean log Θ_*	Standard Deviation
active	107	0.77	0.10
absent	59	1.00	0.09
transition	26	0.90	0.09

TABLE 2

From Table 2 we see a 23 percent reduction in log Θ_* as the jet stream activity changes from absent to active, with the transition day providing an intermediate value.

We note good agreement with microplanometer measurements of Eaton et. al. (1985) who, at the same wavelength, based on 680 measurements at White Sands Missile Range made during periods of strong jet stream activity obtain a mean value of Θ_* of 5.9 urad (log Θ_* = 0.77) while at Capistrano Test Site they obtain, based on 552 measurements made in the absence of such activity, a mean of 8.7 urad(log Θ_* = 0.94).

Eaton et.al.(1985) suggest that the influence of atmospheric waves on C_n^2 be investigated. This is parameter #5 which took on only two values during the times that scintillometer measurements were made. For April 30, May 5 and May 6 (95 profiles, log Θ_* = 0.95) a northerly wind maximum was observed; for May 2-May 4 (92 profiles, log Θ_* = 0.77) a trough. Θ_* appears to be significantly higher for the first binning, which except for the 12 profiles of April 30 corresponds to the absence of the jet stream. The trough is associated with two of the days of jet stream activity and the transition day.

Conclusions

The effects of varying meteorological conditions on isoplanatic angle are statistically significant. Jet stream activity has been seen to reduce this parameter significantly. The influence of atmospheric waves has been noted.

Further study is needed. The effects of combinations of parameters should be investigated. The effects of meteorological conditions on C_n^2 at each level should yield more detailed information on the effects of jet stream activity.

Acknowledgments

We would like to thank the PSU meteorology department for suppling the basic data used in this report. We are grateful to Edmund Murphy of AFGL for sharing the scintillometer data with us. Several very helpful discussions of the statistical aspects of this report were held with Professor Jack Callahan of Massachusetts Maritime Academy.

References

Battles, F.P., E.A. Murphy and J.P. Noonan, "The Contribution of Atmospheric Density to the Drop-off Rate of C_n^2", accepted for Publication, Physica Scripta, 1987.

Born, M., and E. Wolf, Principles of Optics, 6th ed., p. 482, Pergamm Press, Oxford, 1983.

Brockett P. and A. Levine, Statistics and Probability and Their Applications, chapter 9, Saunders College Publishing, 1984.

Eaton, F.D., W.A. Anderson, J.R. Hynes and G. Fernandez, "Isoplanatic Angle Direct Measurements and Associated Atmospheric Conditions", Applied Optics, 24, 3264-3273, 1985.

Hufnagel, R.E., "Propagation Through Atmospheric Turbulence", in The Infrared Handbook , U.S. Government Printing Office, Washington D.C. chapter 6, 1978.

Kukharets, V.P. and L.R. Tsang, "Structure Parameter of the Refractiv Index in the Atmospheric Boundary Layer" , IZV: Atmospheric and Ocean Physics, Vol.16, pp. 397-405, 1980.

Loos, G.C. and C.B. Hogge, "Turbulence of the Upper Atmosphere and Isoplanatism", Applied Optics, Vol. 18, pp. 2654-2661, 1979.

Murphy, E.A., R.E. Good and S. Sheldon, "Stellar Scintillometer Measurements of C_n^2 at Kirtland AFB, New Mexico", AFGL Technical Memorandum , no. 98, 1984.

Murphy, E.A. and F.P. Battles, "Isoplanatic Angle from Stellar Scintillometer Measurements at Pennsylvania State University", AFGL Technical Memorandum, no. 129, 1986.

Ochs, G.R. , Ting-i Wang, R.S. Lawrence , and S.F. Clifford, "Refractive Turbulence Profiles Measured by One-Dimensional Spatial Filtering of Scintillations", Applied Optics, Vol. 15, pp. 2504-2510 1976.

Scheffe, H., The Analysis of Variance, John Wiley and Sons, pp 331-33 1959.

Walters D.L. and K.E. Kunkel, "Atmospheric Transfer Function for Dese and Mountain Locations: The Atmospheric Effects on r.", Journal of th Optical Society of America, Vol. 71, pp. 397-405(1981).

APPENDIX

As indicated in my proposal the $20,000 has been allocated to 4 general areas:

1. Frank P. Battles (Salary)	:	$12000.00
2. Laura L. Kelleher (Salary)	:	$ 2500.00
3. Computer services/equipment	:	$ 3452.59
4. Travel/per diem	:	$ 2047.41
TOTAL	:	$20000.00

Discussion

1. This represents compensation to the principal investigator at a rate of $16.00 per hour for 750 hours.

2. This represents compensation at a rate of $12.50 per hour for 200 hours.

3. The bulk of this expenditure is for the purchase of a Zenith 248 computer system, Zenith EGA color monitor and a Seikosha SL-80AI printer.

4. The bulk of this expenditure is for attendence at professional meetings by myself and other members of the Basic Sciences Department and for travel from the Academy to AFGL.

FINAL REPORT NUMBER 4
REQUESTED A NO-COST TIME EXTENTSION
TO BE SUBMITTED IN 1987 MINI-GRANT FINAL REPORT
Dr. Albert Biggs
760-6MG-072

FINAL REPORT

BASIC RESEARCH ON THE IMPACT OF COGNITIVE STYLES ON DECISION MAKING

Dr. Patricia T. Boggs

John Carroll University
University Hts., OH 44118

Contract No. F49620-85-C-0013/SB5851-0360

INTRODUCTION

In the design and implementation of computerized tools developed to enhance decision making there are an overwhelming number of characteristics of the decision maker and the decision making process which need to be explicitily considered. Yet, it appears as if the driving force behind the development of these tools has been the ability of the hardware and the software with the needs of decision makers having been overlooked. A preferred approach to that of fitting the decision maker to the tools is designing the tools to increase the quality of the decisions. This approach is particularly critical in designing tools for decision making in command and control environments.

This report has two main purposes. The first is to provide background information from the literature related to the characteristics of decision makers that may have an impact on the design of adaptive computer tools. Analysis of decision maker's characteristics, identified in the conceptual and empirical literature, has provided the necessary background for the development and execution of a pilot study. The second purpose of this report is to present the results of the pilot study carried out January, 1987 through October, 1987.

A thorough understanding of this research problem requires studying literature across a number of disciplines. Consequently, studies from management,

management science, management information systems and cognitive psychology literature have been examined. Study of the adaptation of decision enhancing computer tools to specific decision making styles and characteristics of the decision maker also requires an examination of the descriptive, behavioral aspects of decision making.

DATA PROCESSING FOR DECISION MAKING

The dramatic, rapid technological changes in the computing industry led to suboptimal implementation of the tools in environments that had not been studied sufficently. Advances in computing power were confined to creating more and more reports in the business data processing arena which facilitated further misapplications of management science tools by more and more people. In less than thirty years, use of computers progressed from government to big businesses down to the households. Very little analysis has been undertaken to determine the best way for this new technology to facilitate the decision making environment, particularly when decisions were not highly structured. In support of this idea, Mason and Mitroff (1973) contend that:

> the design of MIS has tended to reflect this orientation of their designers, i.e., the designers of MIS have tended to project (or mistake) their dominant psychological type (thinking-sensation) onto that of their clients (p. 478).

A few definitions of decision support systems(DSS) need to be presented. A useful definition by Keen and Scott Morton (1978) is that DSS:

> implies the use of computers to
> 1. Assist managers in their decision processes in semi-structured tasks
> 2. Support, rather than replace, managerial judgment
> 3. Improve the effectiveness of decision making rather than its efficiency (p.1).

The second point needs to be emphasized frequently when dealing with the general public who tend to believe that computers will replace all of us one day. The third point emphasizes that the use of a DSS may, in fact, increase the time required to arrive at a decision but presumably, the decision will be better than without the DSS.

These same concepts are embraced in the definition of DSS developed by Alavi and Henderson (1981):

> ...a DSS is not a computer-based system primarily designed to collect, manipulate, and distribute information. Rather, it is a system linked to the process by which managers arrive at decisions. It's role is not to replace the decision maker, but to enhance his or her effectiveness (p. 1309).

Blanning (1984) refers to DSS as causal model based systems which he distinguishes from expert systems for managers which are judgmental. All of the examples he provides of expert systems for managers appear to be used in situations where the data are easily quantifiable and seem to possess very limited judgment capabilities.

Alter (1980) developed a taxonomy for decision support systems based on generic operations provided by the systems in any functional area or task domain. The operations he

identified and used for classification are:

> Retrieving a single item of information
> Providing a mechanism for ad hoc data analysis
> Providing prespecified aggregations of data in the form of reports
> Estimating the consequences of proposed decisions
> Proposing decisions
> Making decisions (p. 73).

Whatever definition is articulated and whether a system is adaptive or not, it seems clear that to accomplish the goals of a DSS, it must contain data base management software and a very user friendly query capability. Additionally, some type of mathematical modeling should be present (Bonczek, Holsapple, and Whinston, 1981). A relevant research question becomes is there a generic type of decision enhancing tool regardless of decision making style or other characteristics of the decision maker. If not, what attributes impact decision making and how can these be supported in a DSS. These questions are central to this study.

Prior to the computerization of data processing, people processed data primarily manually usually with the assistance of mechanical devices. The methodology utilized to accomplish this processing was easily modeled when computers became available because the tasks were highly structured, were often quantitative or involved classification, sorting or summarizations of some type. The specific characteristics of the people who had previously done the processing manually had little, if any

impact on the design of the software that subsequently accomplished the tasks.

Once the data were processed electronically, other aspects of the organizational functions became candidates for computerization, data processing became more complex and management information systems evolved. Decision making styles and individual management styles and cognitive factors were somewhat more relevant to the design of MIS than they were for the design of EDP systems and yet, no evidence of these considerations in design can be found. Before the computerization of MIS, managers already had their own management information systems to facilitate their operations. These management information systems incorporated more information than could be modeled in the computer and consequently, the resulting computer model was often incomplete. Its primary purpose was to augment the management information system of the manager with regularly recurring reports that would presumably free the managers for more productive activities.

Managers may also have their own, non computer implemented decision support systems. These, however, vary tremendously from one person to another and are used for illstructured or nonstructured problems. The task of modeling these decision support systems is made more difficult because the decision making processes utilized for which modeling is desired are not transparent to either the modelers or even to the decision makers themselves. We simply do not know how decisions are made, what

information is utilized or if any part of the process can be quantified. The significance of cognitive styles cannot be understated at this point in the conceptualization of effective DSS. This is why the study of cognitive psychology, management and management science literature is vital in undertaking DSS analysis.

Management information systems supported the information needs of managers involved in regular, recurring types of activities. The MISs were not particularly suited for aiding those people in unstructured, unique problem domains. Watkins (1980) supports this:

> For the most part, the development of computer based MIS has been directed to the structured and operational control level of decision making. These types of decisions require routine, repetitive, and easily identifiable and retreivable information (p. 6).

DECISION MAKING

Decision Making Phases

Defin'tions of decision making will be explored prior to the examination of decision enhancing tools. Huber (1980) makes a distinction among decision making, choice making and problem solving. Specifically, he defines problem solving as encompassing both choice making and decision making. Choice making occurs after alternative solutions have been generated and includes those activities necessary to evaluate and choose among the alternatives.

Decision making consists of choice making preceded by the activities necessary to identify, define and diagnose the problem and the activities necessary to generate the alternatives which are then evaluated in choice making. Problem solving includes all the components of decision making as well as two steps after the choice is made. These steps are: 1. the activities needed to implement the selected solution and, 2. those necessary to maintain, monitor and review the results of implementing that solution (Huber, 1980, p. 8).

Huber's problem solving definition is similar to Simon's (1977) definition of decision making which "comprises four principal phases: finding occasions for making a decision, finding possible courses of action, choosing among courses of action, and evaluating past choices" (p. 40). Simon labels these phases intelligence, design, choice and review. Huber's second and third categories of decision making are covered in Simon's second phase.

The term used by Hogarth (1980) is decision analysis and incorporates more of the environment into the definition than do others although it stops at choice. Additionally, Hogarth's steps include uncertainty assessment and sensitivity analysis. The steps are listed below:

Step 1: Structuring the problem
Step 2: Assessing consequences
Step 3: Assessing uncertainties
Step 4: Evaluating alternatives
Step 5: Sensitivity analysis

Step 6: Information gathering
Step 7: Choice (Hogarth, 1980 p. 137).

Clearly, Hogarth's approach has an operations research orientation. One of the tasks to be accomplished in step 1 on which Hogarth elaborates is identification of the decision makers which is relevant in the discussion of adaptive decision aids. If an organization does not know who, specifically the decision makers are, then how can the systems be adapted. Although Simon's taxonomy has been used extensively in the literature--usually based on his 1960 work--and Hogarth's has analytical appeal, Huber's definition is preferred by the writer for this discussion. Whatever definition, it is clear that Simon's insight is still valid: "The complexity of decision making has posed grave difficulties in its study and description" (Simon, 1965, p. 36).

Decision Types

The level of decision being made will have an impact on the importance of the various steps in the decision making process. Three levels of decision making are usually identified in the management literature (Davis, 1984). Operational decisions are concerned with the day to day activities of an organization. Tactical decisions usually involve 6 months to 2 years of direct impact on the organization. Strategic decisions have a long term impact and are vital to the viability of the organization. Additionally, each of the steps in decision making is quite

possibly composed of more decision making activities. For example, the implementation of the solution for strategic and tactical levels probably will include decision making junctions.

Another useful way of distinguishing types of decisions is in their frequency of occurence and the procedures available for their solutions. Simon's (1977) nomenclature is used extensively in the management, management science and MIS literature. He describes a programmed or structured decision as one that is regular and recurring and for which unambiguous solution procedures are readily available. A nonprogrammed or nonstructured decision is one which is probably nonrecurring, difficult to formulate, novel and significant to the organization. This distinction is relevant for the design of decision enhancing tools.

ENHANCING DECISION MAKING

Decision making is difficult not only because of the complexity of the task and of the environment, but also because of limitations in human information processing (HIP). Much of the early suggestions for improving decision making were based on comparisons of normative models and observed performance (Isenberg 1984). One of the major limitations in applications of normative decision theory is that problems are not typically structured so parsimoniously to lend themselves to solving using such a

structured format. As Cohen, et al. (1985) have indicated:

> in the case of decision processes, what matters are human goals, perceptions, and emotions that do not lend themselves to quantitative analysis in the spirit of classical mathematics (p. 766).

One could add to this list the point that the decision maker's past experiences and expertise are also critical.

The tradeoffs and ambiguities of normative versus behavioral models are well discussed in Einhorn and Hogarth (1981) who conclude "it is uncertain whether human responses or optimal models are more appropriate" and who "know of no theory or set of principles that would resolve this issue" (p. 59). However, the models provide valuable insight into decision making.

If the purpose of DSS is to enhance decision making then the components which need support must be identified. This is where the dearth of research on performance measures becomes most apparent. Although no clear performance measures have been developed for evaluating decision making, numerous articles have been published offering theories of how decision making can be improved. Much of the work is based on the idea that humans, as processors of information, are rift with biases and limitations that negatively influence their performance in decision making, particularly in complex or stressful situations (Chu and Rouse, 1979). A comprehensive review of these is provided by Remus (1985). He organizes biases into two categories: those associated with presenting data to a decision maker and those associated with the

processing of information. The data presentation biases may occur based on source of information, the format of the data and the sequencing of the data. People tend to filter data selectively based on their past experiences. They retain data with which they have greater familiarity but they incorrectly remember data that is contrary to their expectations. Similarly, they solicit information consistent with their views and will tend to disregard data which contradicts their expectations (Remus, 1985, p. 620).

Information processing biases are sometimes introduced because the heuristics that people employ are biased (Tversky and Kahneman, 1974). Other problems with heuristics highlighted by Remus include limiting solution domain to previously encountered situations and inconsistency in the use of heuristics. Additionally, people consistently misinterpret and misapply statistical techniques such as making inferences from nonrepresentative small samples and ignoring uncertainty. They tend to prematurely restrict search strategies and conservatively revise probabilities in light of new information (Remus, 1985, pp.621-622).

Hogarth (1980) and Hogarth and Makridakis (1981) present a concise summary of HIP biases organized by what they have identified as different stages of information processing--information acquisition, information processing, output and feedback. Apparently the summary by Remus is influenced by these earlier works. The categories

utilized by Hogarth and Hogarth and Makridakis provide useful distinctions for the design of tools.

One of the shortcomings of HIP identified is that people prematurely bound the number of alternatives (Remus, 1985; Shaklee and Fischoff, 1979; Payne, 1976). Yet Goslar, Green and Hughes (1986) reached a contrary conclusion in a recent study. In determining whether significant differences exist in the number of alternatives considered by subjects evaluating a marketing case by those with decision support systems available versus those without, they found that subjects who were not given access to DSS considered more alternatives. This result is contrary to what would be expected from implementation of a decision enhancing tool. Additionally, the authors found no significant impact on subjects' confidence in decision making, the processes employed or performance levels for whom the DSS was available. The authors point to a number of experimental influences which may have contributed to these somewhat contradictory results. Their study did, however, recognize the importance on nonambiguous performance measures.

These limitations in HIP may be offset by the development and implementation of appropriate decision support tools (Remus, 1985; Zmud, 1979). The basic idea is that humans tend to suffer the aforementioned processing deficiencies and can, therefore, benefit from an association with unbiased decision enhancing tools that make up for the human deficiencies. To determine the

extent of the forthcoming benefits, empirical analyses with unambiguous measures of performance are required.

PSYCHOLOGICAL CHARACTERISTICS AND DECISION MAKING

Cognitive style research has for the most part been based on the Jungian view of perception and evaluation. Cognitive style and decision style are used synonomously in the literature. Categorization of a person's cognitive style is generally accomplished with the Myers-Briggs Type Indicator personality assessment inventory (Myers, 1977). These two, assumed independent dimensions (perception and evaluation) result in four distinct nonordered styles--sensation-feeling, sensation-intuition, intuition-thinking, and intuition-feeling. Keen and Bronsema (1981) provide the following explanations of these human uses of information:

> 1. "Finding Out" [perception]
> Sensing: preference for known facts: reliance on concrete data and experience
>
> Intuition: looking for possibilities and relationships; focus on concepts and theory
>
> 2. "Deciding" [evaluation]
> Thinking: judgments are based on impersonal analysis and logic
>
> Feeling: judgments are based on feelings and personal values (p. 31).

Because Jung was concerned with how people absorb the information and how they use it, a natural extension of his theory is to elaborate on the different ways of presenting

information which would tend to complement different cognitive types. This elaboration is first articulated by Mason and Mitroff (1973).

Two other dimensions defined by Jung, introversion (extroversion) and judging (perceiving) when combined with the information dimensions form sixteen distinct types as measured by the MBTI.

A study by Alavi and Henderson (1981) used the Jungian typology to study the effects of decision styles on decision support system utilization and on user evaluations of the system. Their findings supported the contention in the literature (Doktor and Hamilton, 1973; Mason and Mitroff, 1973; McKenney and Keen, 1974) that individual learning processes should be explicitly considered in the design and implementation of a computer information tools. Specifically, their "results indicate decision style as conceptualized by the Jungian typology will influence utilization of a model based information system" (Alavi and Henderson, 1981). Differences in cognitive style are credited as being responsible for explaining differences in the way managers and management scientists think (McKenney and Keen, 1986) and in their behaviors (Henderson and Nutt, 1980). Although there appears to be general agreement in the literature that cognitive style should make a difference in the utilization of information, the form of presentation preferred and consequently the quality of decisions, empirical validation is simply not evident.

CONCLUSIONS FROM THE LITERATURE

Clearly, the literature suggests a need for more careful analysis in the conceptualization of adaptive decision support systems. Adaptation could take the form of altering the presentation of the material (such as spatial vs. semantic) or of altering the type of material presented such as regression results, expert opinions or other methods of overcoming some HIP limitations.

It may be that the crucial role of a decision support system would be to support the particular type of decision making style required by the task. Thus, the adaptation may be primarily task dependent and then secondarily modified for different decisionmaking styles. Particularly in decision making situations where the temporal demands are severe, having a decision support system provide the proper type of information for the specific type of task at hand would possibly enable the decision maker to focus on the problem at hand.

If the purpose of a decision support system is to enhance decision making (or enhance organizational structure or make it more conducive to decision making) then one must first determine which components need support.

As many authors have recognized, it is necessary to develop a body of scientific research in decision making with nonambiguous measures of performance whose external

validity can be supported. There have not been systematic scientific procedures followed in past research to allow meaningful conclusions to be drawn regarding adaptation of decision support systems. Once relationships can be established with exploratory research between cognitive styles, expertise, HIP weaknesses, these relationships can be used to determine the extent of their impact on decision performance and thus the contribution of adaptive decision support systems to decision making can be ascertained.

PILOT STUDY

As a first step in the background research in determining the feasibility of adaptive DSS, it is necessary to determine if cognitive style impacts performance in decision making. As researchers have suggested as part of their conceptualization of decision making paradigms that "decision aids are provided to overcome limitations inherent in human cognition" (Zmud, 1979, p. 968) then establishing empirical support is necessary before the investment of resources in such aids is undertaken.

This pilot study was designed to accomplish three specific goals. The first goal was to provide an exploratory analysis of the relationship of cognitive style to performance in decision making. Because the study is exploratory, no formal hypotheses are stated a priori although the appropriate two sided null would be that

cognitive style does not have a significant impact on decision making performance. Decision making performance was measured by results obtained for each subject's decisions in a management simulation game called Tempomatic.

The second goal was to provide feedback on the appropriateness of using a decision making simulation for this type of research. As discussed in the literature review, one of the main problems in analyzing decision making is the lack of performance measures with levels of measurement higher than ordinal data. Therefore, whether this methodology could be implemented effectively was considered an important outcome of the study.

Another goal was to determine the validity of using business students as subjects in a study to draw inferences about command and control personnel. It was postulated that a sufficient number of similarities in decision making environments and in decision makers exist between management decision making and command and control decision making to allow generalizations to be made between the two groups. The major difference in the problem domains is the temporal component of decision making in command and control. Because the number of command and control subjects in this pilot study was anticipated to be small, qualitative comparisons between command and control subjects and business student subjects were originally planned to be undertaken to support the external validity

of using business students as subjects in decision making experiments.

DATA COLLECTION AND RESULTS OF THE PILOT STUDY

The subjects for this study came from two sources: 1. volunteer Air Force personnel at Wright Patterson Air Force Base in Dayton, Ohio and 2. volunteers in the MBA program at Wright State University in Dayton, Ohio. The Air Force personnel at WPAFB who were identified as having been fighter pilots or navigators at some point in their Air Force careers were sent a letter asking them to participate in the study. A reply form was included with the letter with instructions to return to AFHRL/LRG indicating their participation decision. Nineteen officers responded affirmatively to the request.

MBA students at Wright State University were contacted via personal letters delivered to their campus addresses. A reply form was included with their letters with instructions to return to the investigator at the campus address. Twenty one MBA students responded affirmatively to the request.

In both groups, those who were interested in participating were asked to provide their telephone numbers to facilitate scheduling of briefing sessions. Three different sessions of two hours were scheduled for each group of participants. All WPAFB sessions were held at AFHRL/LRG; all WSU sessions were held in the College of

Business and Administration at Wright State University. At this two hour orientation and training session, the purpose of the study was conveyed to the participants, the research team was introduced, and the rules of the Tempomatic simulation were explained and interpretation of the output was given. The research team consisted of the principal investigator, her graduate assistant provided by the Department of Management at Wright State University, an MBA level student worker who was an expert in the implementation and administration of the Tempomatic simulation, and an Air Force Major who was an expert in the use, administration and interpretation of the MBTI. A total of seventeen Air Force volunteers came to the sessions and eighteen MBA students attended.

The Myers Briggs Type Indicator for cognitive style classification also was administered at the initial briefing session. Keen and Bronsema (1981) provide a number of convincing arguments in support of the Myers Briggs Type Indicator in cognitive style research:

1. the Jungian paradigm is theoretically strong
2. it is reliable and well designed
3. it has been used extensively and, hence is backed by many years of data
4. it is a good discriminator for behavior relevant to MIS (p. 24).

Materials the volunteers needed to participate were also distributed at this session. Each participant was

given a notebook, background information on the hypothetical company, summaries of the rules of the simulation, decision sheets for recording their decisions and information sheets with each member of the research team's affiliation and telephone number listed. Samples of the handouts are included in Appendix A. Consent forms for both the Air Force and Wright State University were also signed by the participants at this meeting.

At the conclusion of the initial briefing session, each of the volunteers was asked to complete a decision sheet for a practice quarter and to submit that to the research team for processing. The simulation was executed on the IBM 3083 at Wright State University. After the results were determined, each participant was contacted to meet with the principal investigator to go over the results to be sure each understood how the game worked and to provide an opportunity to have questions answered. The results from the practice quarter were not used in the study.

Training of subjects and administering the game to both the command and control subjects and the MBA volunteers involved an extensive amount of preparation and supervision. To avoid confounding the task by having subjects enter decisions themselves to the computer they gave their decisions to the principal investigator or one of the research assistants who submitted all input to the IBM for processing.

Participants were given a number of choices regarding how to get their decisions to the research team for computer processing. All had the choice of calling in their choices to one of three people. The Air Force volunteers also had the option of dropping their responses off to the AFHRL/LRG office; mailing via base mail or having someone pick up at their offices. The same choices were available for delivery of computer output after each decision was processed. Because the MBA volunteers did not all work at one facility, their choices were somewhat more limited. They could choose to drop off/pick up at the principal investigator's office door in the Business Building at Wright State University or they could use the campus mail and student mailbox combination. Some of the MBA volunteers were also Air Force officers and therefore had the full range of choices.

Not surprisingly, some of the subjects who began the experiment chose not to finish. Some of those who dropped out indicated to me they had not anticipated the study would require so much work. Others simply quit with no explanation. Table I summarizes the participants by affiliation.

TABLE I

DISTRIBUTION OF SUBJECTS BY AFFILIATION AND EXTENT OF COMPLETION OF EXPERIMENT

	AF	MBA
Responded "yes" to initial letter	19	21
Attended briefing & completed MBTI	17	18
Started simulation	14	15
Finished experiment	6	10

A participant finished the experiment if he/she completed the simulation for eight quarters. The performance measure was obtained by summing the net income for each of participants over the eight quarters of simulated business activity.

One of the goals of the pilot study was to relate cognitive style to performance. Table II shows the distribution of MBTI categories by affiliation of participant.

TABLE II

MBTI CATEGORY BY AFFILIATION

Category	AF	MBA
ENTJ	2	2
ENTP	2	0
ESFJ	0	1
ESTJ	5	1
INTJ	1	2
INTP	2	3
ISFJ	0	3
ISTJ	5	4
ISTP	0	2

The rankings and MBTI category are supplied in Tables III, IV and V for each group separately and the combined group of participants.

TABLE III

RANK AND MBTI CATEGORY FOR AIR FORCE SUBJECTS

Rank	MBTI
1	ISTJ
2	ISTJ
3	ISTJ
4	ESTJ
5	ENTJ
6	ENTP

TABLE IV

RANK AND MBTI CATEGORY FOR MBA SUBJECTS

Rank	MBTI
1	ENTJ
2	ISFJ
3	ISTJ
4	ISFJ
5	ESFJ
6	INTJ
7	ENTJ
8	ISTJ
9	ISTJ
10	ISTJ

TABLE V

COMBINED RANK AND MBTI CATEGORY

Rank	Affiliation	MBTI
1	MBA	ENTJ
2	MBA	ISFJ
3	MBA	ISTJ
4	MBA	ISFJ
5	MBA	ESFJ
6	MBA	INTJ
7	AF	ISTJ
8	AF	ISTJ
9	MBA	ENTJ
10	AF	ISTJ
11	MBA	ISTJ
12	MBA	ISTJ
13	AF	ESTJ
14	AF	ENTJ
15	AF	ENTP
16	MBA	ISTJ

Unfortunately, the number of subjects completing the experiment was too small to allow any multivariate, parametric test of significance to be conducted. Descriptive discussion is the extent of the appropriate statistical analysis.

The subjects receiving the highest rankings of performance (the top six finishers) were MBA volunteers. The simulation game used in the study to measure decision making performance was developed for business students taking policy courses--the capstone course for undergraduate business majors. Most of the AF volunteers had engineering undergraduate backgrounds. However, a chi square test for independence of classification performed on the data organized as in Table VI revealed no significant difference between AF and MBA subjects' performances.

TABLE VI

Distribution of Subjects by Performance and Affiliation

	Performance	
Affiliation	Top Half	Bottom Half
AF	2	4
MBA	6	4

The MBA volunteers also completed the task more quickly than most of the AF volunteers. In fact, one of the AF

volunteers did not complete the simulation until September, 1988. All of the AF subjects called the principal investigator with questions regarding the simulation at least one time during the experiment. Only three of the MBA subjects contacted the investigator with simulation questions. Clearly, evaluating performance in decision making with existing business simulations is not appropriate with command and control personnel. To effectively evaluate decision making performance, a problem solving domain familiar to the subjects incorporating more of the components of command and control decision making would be preferred to easily obtained, well documented business simulations. Although a good deal of evidence exists to support the use of simulations in business and information research, the approach was very difficult to implement in this defense research. This conclusion is based not only on performance of AF personnel but on the logistical problems and other implementation problems experienced by the research team and noted differences between the two groups of subjects.

For many of the same reasons noted above, this study cannot support the notion that results of studies utilizing business school subjects should be generalized to be applicable to defense environments. The business students in studies utilizing business scenarios have an implicit advantage over non business subjects participating simply because of the background breadth--something which cannot be conveyed in training sessions for experiments.

Another objective of this pilot study was to determine if performance in decision making was related to cognitive style. The results of the study do not support any relationship between cognitive style and decision making performance. In the group of Air Force subjects, the top three performers were all ISTJs; in the group of MBA subjects, the bottom three performers were all ISTJs.

Only 9 of the 16 MBTI categories were represented in the total of volunteers beginning the experiment; only 7 of the 16 MBTI categories were represented in subjects completing the study. While some studies use only two of the four dimensions, resulting in 4 total categories versus 16, the sample size resulting here would still be too small for meaningful statistical applications.

In summary, with respect to the objectives of the pilot study, the number of subjects was too small to complete any parametric tests of significance on the data with respect to performance in complex deicision making environments. Additionally, using a business tactical decision making simulation with command and control personnel may have confounded the results even if the sample size had been sufficient for further analysis.

The main feedback on the logistics of using a simulation involving such a large investment of time by both the subjects and the research team is that both groups in this study were at times overwhelmed by the complexity of the task. The subjects invested a great deal of time in

making their decisions. Two members of the research team also had to make the same number of decisions with each of the subjects as their decisions provided the competition for each of the subjects in the study. In addition to the same amount of work as the subjects in terms of decision making, the principal investigator and graduate assistant also had to input all the data to the computer for processing, acquire and distribute the output and obtain new decisions from the subjects. The process was simply too complex for the number of people on the team. To effectively conduct future research such as this, more resources, particularly in terms of computer support would have to be allocated to the design.

Finally, a caveat with respect to basing future research on past empirical studies utilizing convenience samples is stated based on this experience. No studies exist to support the validity of drawing conclusions for a specific subset--command and control decision makers-- of the population from research conducted on these convenience samples. If reliable, unbiased research is be conducted that is unambiguously relevant for command and control personnel decision making, resources will have to be allocated that allow that particular group to be sampled using problem solving domains that are appropriate in that environment.

Appendices can be obtained from

Universal Energy Systems, Inc.

FINAL REPORT FOR THE 1986-87 MINI GRANT PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

For

A Feasibility Study and Test Application of Uncertainty Estimates to an Atmospheric Dispersion Model with Potential Utility in Air Force Operations

Submitted by: Thomas A. Carney
Assistant Professor
Department of Meteorology,
Florida State University
(904) 644-1268

1986 SFRP Location: Air Force Engineering and Services
Center, Environics Division,
Tyndall AFB, FL 32403-6001

Period of Research: 1 January 1987 to 31 December 1987

1. INTRODUCTION

During the summer of 1986, the author participated in the AFOSR/UES Summer Faculty Fellowship Program at the Air Force Engineering and Services Center (AFESC), Tyndall AFB, Florida. In consultation with Capt. Larry E. Key, several models of atmospheric dispersion were considered for potential use simulating releases of toxic or hazardous chemicals. The research reported here was undertaken under a research initiation grant awarded to study the feasibility of implementing specific improvements to this class of models.

It has long been recognized that dispersion models would be more useful to decision makers if the uncertainty associated with model predictions were quantitatively available. This information could then be used as an additional element of the decision matrix to improve overall confidence in decisions made based on modeling results. The uncertainty associated with a particular model estimate may be divided into:

1) model formulation uncertainty, resulting from inadequate description of the modeled event by the simplified model parameterizations;
2) stochastic uncertainty due to the random nature of the atmospheric processes responsible for transport and dispersion; and
3) measurement uncertainty, resulting from inherent imprecision or inaccuracy of meteorological measurements.

For the convenience of this study, the first two of these will be further grouped together as "model" uncertainty since they both represent possible shortcomings in the model's representation of the actual physics involved in the dispersion process. In general, model uncertainty is believed to be much greater than measurement uncertainty and some estimates

of uncertainties in model results based on the departure of the modeled event from the simplifying assumptions have been derived (e.g., Turner,1970; Pasquill,1974). In addition, there has been considerable emphasis on quantifying model uncertainty in many recent studies. Comparatively few attempts to quantify uncertainty due to measurements have been made, however, it may be possible to infer such estimates from some model sensitivity studies. Freeman *et al.* (1986) referenced a number of pertinent model sensitivity analyses and proposed a method of incorporating uncertainties due to measurements into dispersion models.

During the course of this feasibility study, a decision was made to concentrate on the propagation of measurement error in dispersion models. This decision was partially influenced by the award of an Air Force small-business grant to Sigma Research Corporation to study hazard response modeling uncertainty, a study that begin in July, 1987. Subsequent meetings with the principle investigator of that study have indicated that focusing the present study on the measurement error would be the most complementary and least redundant approach.

Based on the evaluation of models performed during the 1986 Summer Faculty Fellowship at AFESC, the dispersion model used in this project is the Air Force Geophysical Laboratory (AFGL) AFTOX model. AFTOX is an acronym for the Air Force Toxic Chemical Dispersion Model which is currently under development at AFGL. AFTOX is intended to provide general dispersion modeling capabilities to field units throughout the Air Force. Consequently, users of the model may be far removed from the considerations involved in its development; thus, quantification of the propagation of uncertainty in this model should be particularly valuable.

Much of the previous work concerned with the propagation of errors in dispersion models has focused on Gaussian-plume models. These models are

most appropriate for continuous elevated sources, however, chemical spills have typically been more successfully modeled with "puff" formulations. The puff configuration is a closer representation of instantaneous releases or spills which have a short release time relative to the time of travel to points of interest. Puff models may also be used to represent continous releases by considering a series of puffs to simulate a continous plume. Typically, a number of puffs will contribute to the concentration at a particular point in such a scenerio.

The difference in the mathematical representation of a Gaussian-plume and a Gaussian-puff may be illustrated by comparing the following equation, from AFTOX, representing a puff configuration, to the standard Gaussian-plume equation (*e.g.*,Freeman *et al.*, 1986; Hanna *et al.*, 1982).

$$G(x,y,z,t-t') = \frac{Q(t')}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z} \cdot \exp\left[-\frac{1}{2}\left[\frac{x-u(t-t')}{\sigma_x}\right]^2\right] \cdot$$

$$\exp\left[-\frac{1}{2}\left[\frac{y}{\sigma_y}\right]^2\right] \cdot \left[\exp\left[-\frac{1}{2}\left(\frac{z-H}{\sigma_z}\right)^2\right] + \exp\left[-\frac{1}{2}\left(\frac{z+H}{\sigma_z}\right)^2\right]\right]$$

where,

t′- time of emission.

(t-t′) = elapsed time since emission

Q - concentration of the puff at emission.

u - wind speed at a height of 10m.

σ_x- standard deviation of Q in the x direction.

σ_y- standard deviation of Q in the y direction.

σ_z- standard deviation of Q in the z direction.

x - the downwind axis.

y - the horizontal axis perpendicular to the plume axis, with y = 0 at the plume axis.

z - height above the ground (ground is assumed to be uniform).

H - effective height above ground level.

Note that the spread parameters (sigmas) commonly used in these equations were empirically derived from experiments generally appropriate for continuous plume considerations. There have been very few experiments to determine dispersion parameters for "puffs". As a consequence, when the standard Pasquill-Gifford parameters are used with puff models, an additional element of uncertainty is introduced. This uncertainty might properly be regarded as "model" uncertainty and outside the focus of this study, however, it will be useful to consider how measurement errors are propagated differently by the respective formulations. A key element of the present analysis of the propagation of uncertainty in AFTOX is the comparison of these results with previous analyses conducted for Gaussian-plume models.

2. METHODOLOGY

The following equation for the propagation of error after Freeman *et al.* (1986) was used to examine the uncertainty in computed concentrations from AFTOX associated with specified uncertainties in input variables.

$$S_G^2 = \sum_{i=1}^{n} \left[\frac{\delta G}{\delta x_i}\right]^2 S_{x_i}^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} \left[\frac{\delta^2 G}{\delta x_i \delta x_j}\right]^2 S_{x_i}^2 S_{x_j}^2$$

$$+ 1/2 \sum_{i=1}^{n} \left[\frac{\delta^2 G}{\delta x_i^2}\right]^2 S_{x_i}^4$$

where x_i and x_j repeset the variables assumed to have known uncertainties and the "S"s represent standard deviations determined by the specified uncertainty in the variables. The partial derivatives of this equation are provided in Appendix A.

The AFTOX variables assumed to have known uncertainties are:

a) the source strength, Q;

b) the wind speed, u;

c) the effective source height, H; and

d) the Gaussian spread parameters, sigma-x, sigma-y and sigma-z.

Wind speed is likely to be routinely measured at potential sites for AFTOX applications and current calibration specifications for Air Force instrumentation can provide an estimate of the inherent error associated with such measurements. The wind speed, u, of the dispersion equation is, more specifically, the representative transport wind. AFTOX has routines to adjust the input wind speed to be more representative of the transport wind if, for example, the available measurement is at an inappropriate height. There will be additional uncertainty introduced by such routines, but the assumption made for the purpose of this study is that the total uncertainty

in the input parameters considered, wind speed in this case, is known and can thus be specified *a priori*.

This assumption is particularly important for the other specified parameters which, in general, are not directly measured. AFTOX has an evaporation module which may be used to derive the source strength from other details of a spill or leak. The sigma values are derived from Pasquill-Gifford curves after the stability is either estimated from descriptions of related meteorological conditions or explicitly specified. And, the effective source height for buoyant emissions can be derived from stack and plume characteristics using a plume rise algorithm.

For this study, the uncertainty in these parameters has been specified as 10 or 20 percent of the input value. The values for the parameters themselves were specified in order to avoid the model algorithms designed to calculate or adjust them. Downwind and crosswind distances were assumed to be precisely known, in contrast to the earlier study by Freeman *et al.* (1986) where relative distance was parameterized in terms of wind direction, which was then assigned an uncertainty. The uncertainty values, 10 and 20 percent, were chosen arbitrarily but are thought to represent reasonable error for these parameters, if only "measurement" errors are considered.

The assumptions inherent to the propagation formula are:

a) All measurement inputs except for the above indicated parameters are known exactly;

b) The errors in the specified parameters are random, uncorrelated, and normally-distributed with means equal to zero and standard deviations as indicated in the propagation equation; and

c) The third and higher terms of the Taylor expansion of G around the true values of the specified parameters are negligible with

respect to the first and second order terms.

Of the three, the second assumption is likely the least representative for the current application. The lower boundary of these parameters is zero so a skewed distribution may be more representative. Irwin *et al.* (1987) suggests that a log-normal distribution is more appropriate. In addition, the uncertainty among some the specified parameters might well be correlated. As indicated above, values for the indicated parameters are specified for these models runs, so they are not determined by the model from other input variables. Since the wind speed is used in the model algorithms to determine the effective stack height, the evaporative source strength, and, indirectly, the Gaussian spread parameters, errors in these parameters might otherwise be correlated to errors in this common factor.

Model Runs

Tables 1 and 2 list the conditions used for the model runs in this study. Table 1 conditions were used for simulations with emission releases that could be considered instantaneous and repesented by a single puff. In Table 2, a release that occurs over a finite time is described. Such a release would be represented in AFTOX by a series of puffs, 4 per minute.

Three stability classes (Pasquill types A, D and E) were studied and concentrations at elevations of 1 foot and 10 feet above the surface were derived from the model for each set of conditions. Runs were also made with a prescribed low-level inversion (150 ft) to investigate the effect of limited vertical dispersion.

Table 1: *Environmental variables used for the "instantaneous release" model runs (a single puff).*

TEMPERATURE: 70 F WIND
DIRECTION: 270
WIND SPEED: 7 kts
CLOUD COVER: 1/8
CLOUD TYPE: LOW (St, Ns, fog)
SPILL SITE ROUGHNESS LENGTH: 10 cm
AMOUNT SPILLED: 50 gal
ELAPSED TIME SINCE START OF SPILL: 10 minutes
MAXIMUM CONCENTRATION AT 330 FT
CROSSWIND DISTANCE: 165FT for 1650 & 3300, 330FT for 8250

Table 2: *Environmental variables used for the "continuous release" model runs (20 puffs - 4/min).*

TEMPERATURE: 15 C
WIND DIRECTION: 270
WIND SPEED: 3.6 m/s
CLOUD COVER: 7/8
CLOUD TYPE: LOW (St, Ns, fog)
SPILL SITE ROUGHNESS LENGTH: .5 cm
EMISSION RATE: 1044 kg/min
HEIGHT OF LEAK ABOVE THE GROUND: 2.5 m
DURATION OF SPILL: 5 min
ELAPSED TIME SINCE START OF SPILL: 10 min
MAXIMUM CONCENTRATION AT 1113 m
CROSSWIND DISTANCE: 50m for 1000 & 1750, 100m for 2500

3. RESULTS

Tables 3 and 4 contain the calculated concentrations in grams per cubic meter for the point of the maximum concentration and three downwind distances at elevations of 1 and 10 ft, respectively. Table 5 contains the comparable calculated concentrations for the inversion cases. The model results shown in these tables are for releases specified as instantaneous and Table 6 shows model-derived concentrations for the continuous release scenerios.

The uncertainty in the model results given by the propagation formulation for specified uncertainties in the input parameters is shown in Tables 7-14.

The instantaneous-release runs simulated non-buoyant ground-level spills, consequently, effective source height, H, did not contribute uncertainty to the calculated concentrations. The major contributor to uncertainty appears to be the wind speed for all of these cases, however, at the greatest distance shown (8250 ft) the effect of the prescribed limit to vertical dispersion increases the significance of sigma-z for the inversion cases. At this distance, however, the calculated concentrations are very low so the standard deviations may not be reliable indicators of the uncertainty.

The releases with a finite duration yielded similar results for the cases with no prescribed limit to vertical dispersion. Several inversion heights were modeled for this emission scenario but no discernible effect on calculated concentrations resulted, so those runs are not summarized here.

Table 3: *Model-derived concentrations (g/m^3) at an elevation of 1 foot, 10 minutes after an instantaneous surface release.*

stability	330ft	1650ft	3300ft	8250ft
4.5	0.167	4.7E-3	4.6E-3	6.5E-7
3.5	0.173	5.0E-3	2.7E-3	8.6E-6
1.5	3.2E-2	1.4E-3	4.0E-4	9.3E-6

Table 4: *Model-derived concentrations (g/m^3) at an elevation of 10 feet, 10 minutes after an instantaneous surface release.*

stability	330ft	1650ft	3300ft	8250ft
4.5	0.147	4.6E-3	4.5E-3	6.4E-7
3.5	0.150	5.0E-3	2.6E-3	8.5E-6
1.5	3.1E-2	1.4E-3	4.0E-4	9.3E-6

Table 5: *Model-derived concentrations (g/m^3) at an elevation of 1 foot, 10 minutes after an instantaneous surface release. There is an inversion present 150 feet above the surface.*

stability	330ft	1650ft	3300ft	8250ft
4.5	0.167	4.7E-3	4.6E-3	6.9E-7
3.5	0.174	5.0E-3	2.7E-3	1.3E-5
1.5	3.2E-2	2.0E-3	1.2E-3	7.1E-5

Table 6: *Model-derived concentrations (g/m^3) at an elevation of 1 m, 10 minutes after a continuous release from a height of 2.5 m.*

stability	1000m	1113m	1750m	2500m
4.5	2.3E-4	3.2	0.94	2.9E-7
3.5	1.0E-2	1.4	0.56	1.6E-4
1.5	2.3E-2	0.15	0.07	1.9E-3

Table 7: *Calculated concentration uncertainties (in g/m^3) due to a 10% uncertainty in individual measured variables and all measured variables combined. This is an instantaneous surface release measured at an elevation of 1 foot.*

	U	Q	σ_x	σ_y	σ_z	ALL
STAB: 4.5						
<100 m	0.214	0.017	0.010	0.017	0.017	0.222
1650 ft	4.93E-3	2.37E-4	1.95E-4	4.64E-4	2.39E-4	5.06E-3
3300 ft	2.83E-3	1.23E-4	9.34E-5	2.27E-5	1.24E-4	2.88E-3
8250 ft	1.04E-5	4.58E-8	5.77E-7	1.59E-8	4.62E-8	1.05E-5
STAB: 3.5						
330 ft	0.119	0.017	0.010	0.018	0.018	0.128
1650 ft	1.89E-3	1.72E-4	1.27E-4	7.19E-5	1.73E-4	1.95E-3
3300 ft	5.97E-4	4.87E-5	3.85E-5	2.95E-5	4.92E-5	6.17E-4
8250 ft	1.69E-5	3.23E-7	1.60E-6	2.25E-7	3.26E-7	1.72E-5
STAB: 1.5						
330 ft	0.010	1.64E-3	1.05E-3	1.66E-3	1.66E-3	0.011
1650 ft	9.76E-5	2.33E-5	1.75E-5	1.52E-5	2.35E-5	1.11E-4
3300 ft	1.71E-5	3.72E-6	2.87E-6	3.37E-6	3.75E-6	1.93E-5
8250 ft	9.97E-7	1.23E-7	1.37E-7	1.14E-7	1.24E-7	1.08E-6

Table 8: *Calculated concentration uncertainties (in g/m^3) due to a 20 % uncertainty in individual measured variables and all measured variables combined. This is an instantaneous surface release measured at an elevation of 1 foot.*

	U	Q	σ_x	σ_y	σ_z	ALL
STAB: 4.5						
330 ft	0.679	0.034	0.021	0.035	0.035	0.709
1650 ft	0.017	4.74E-4	4.03E-4	9.57E-4	4.92E-4	0.018
3300 ft	9.66E-3	2.46E-4	1.96E-4	6.47E-5	2.56E-4	9.94E-3
8250 ft	4.07E-5	9.16E-8	1.51E-6	3.32E-8	9.51E-8	4.08E-5
STAB: 3.5						
330 ft	0.317	0.035	0.020	0.036	0.036	0.350
1650 ft	5.93E-3	3.43E-4	2.62E-4	1.90E-4	3.57E-4	6.23E-3
3300 ft	1.92E-3	9.74E-5	7.99E-5	5.91E-5	1.01E-4	2.01E-3
8250 ft	6.02E-5	6.46E-7	3.28E-6	4.52E-7	6.71E-7	6.14E-5
STAB: 1.5						
330 ft	0.022	3.29E-3	2.52E-3	3.42E-3	3.41E-3	0.027
1650 ft	2.61E-4	4.65E-5	3.68E-5	3.04E-5	4.84E-5	2.98E-4
3300 ft	4.66E-5	7.44E-6	6.02E-6	6.87E-6	7.72E-6	5.28E-5
8250 ft	2.22E-6	2.45E-7	3.00E-7	2.33E-7	2.55E-7	2.62E-6

Table 9: *Calculated concentration uncertainties (in g/m^3) due to a* 20% uncertainty *in individual measured variables and all measured variables combined. This is an instantaneous surface release measured at an* elevation of 10 feet.

	U	Q	σ_x	σ_y	σ_z	ALL
STAB: 4.5						
330 ft	0.594	0.029	0.018	0.031	0.022	0.620
1650 ft	0.017	4.60E-4	3.91E-4	9.27E-4	4.44E-4	0.017
3300 ft	9.54E-3	2.43E-4	1.94E-4	6.40E-5	2.46E-4	9.81E-3
8250 ft	4.05E-5	9.12E-8	1.50E-6	3.31E-8	9.40E-8	4.07E-5
STAB: 3.5						
330 ft	0.274	0.030	0.018	0.031	0.022	0.301
1650 ft	5.84E-3	3.38E-4	2.58E-4	1.87E-4	3.39E-4	6.13E-3
3300 ft	1.91E-3	9.69E-5	7.95E-5	5.88E-5	9.94E-5	2.00E-3
8250 ft	6.00E-5	6.45E-7	3.28E-6	4.52E-7	6.68E-7	6.13E-5
STAB: 1.5						
330 ft	0.021	3.14E-3	2.41E-3	3.27E-3	2.94E-3	0.025
1650 ft			% UNCERTAINTIES ARE			
3300 ft			THE SAME			
8250 ft			AS THE 1FT LEVEL			

Table 10: *Calculated concentration uncertainties (in g/m^3) due to a* 10% uncertainty *in individual measured variables and all measured variables combined. This is an instantaneous surface release measured at an elevation of 1 foot. There is an* inversion *present at* 150 feet *above the surface.*

	U	Q	σ_x	σ_y	σ_z	ALL
STAB: 4.5						
330 ft	0.214	0.017	0.010	0.017	0.017	0.222
1650 ft	4.93E-3	2.37E-4	1.95E-4	4.64E-4	2.39E-4	5.06E-3
3300 ft	2.83E-3	1.23E-4	9.35E-5	2.27E-5	1.24E-4	2.88E-3
8250 ft	1.12E-5	4.90E-8	6.18E-7	1.70E-8	7.74E-5	7.83E-5
STAB: 3.5						
330 ft	0.119	0.017	0.010	0.018	0.018	0.128
1650 ft	1.88E-3	1.72E-4	1.27E-4	7.19E-5	1.73E-4	1.95E-3
3300 ft	6.04E-4	4.92E-5	3.89E-5	2.98E-5	2.56E-5	6.22E-4
8250 ft	2.55E-5	4.87E-7	2.41E-6	3.40E-7	9.92E-5	1.03E-4
STAB: 1.5						
330 ft	0.010	1.64E-3	1.05E-3	1.66E-3	1.66E-3	0.011
1650 ft	1.44E-4	3.42E-5	2.57E-5	2.23E-5	6.72E-5	1.74E-4
3300 ft	5.13E-5	1.11E-5	8.61E-6	1.01E-5	4.00E-5	7.00E-5
8250 ft	7.67E-6	9.44E-7	1.05E-6	8.78E-7	1.90E-5	2.09E-5

Table 11: *Calculated concentration uncertainties (in g/m^3) due to a* 10% uncertainty *in individual measured variables and all measured variables combined. This is an instantaneous surface release measured at an elevation of* 10 feet. *There is an* inversion *present at* 150 feet.

	U	Q	σ_x	σ_y	σ_z	ALL
STAB: 4.5						
330 ft	0.187	0.015	9.01E-3	0.015	0.011	0.194
1650 ft	4.78E-3	2.30E-4	1.88E-4	4.49E-4	2.17E-4	4.90E-3
3300 ft	2.79E-3	1.22E-4	9.22E-5	2.24E-5	1.19E-4	2.84E-3
8250 ft	1.11E-5	4.89E-8	6.17E-7	1.70E-8	7.79E-5	7.88E-5
STAB: 3.5						
330 ft	0.103	0.015	8.82E-3	0.015	0.011	0.109
1650 ft	1.86E-3	1.69E-4	1.25E-4	7.08E-5	1.65E-4	1.92E-3
3300 ft	6.00E-4	4.89E-5	3.87E-5	2.97E-5	2.44E-5	6.19E-4
8250 ft	2.55E-5	4.88E-7	2.41E-6	3.40E-7	9.91E-5	1.02E-4
STAB: 1.5						
330 ft	9.75E-3	1.57E-3	1.01E-3	1.59E-3	1.44E-3	0.011
1650 ft			% UNCERTAINTIES ARE			
3300 ft			THE SAME			
8250 ft			AS THE 1FT LEVEL			

Table 12: *Calculated concentration uncertainties (in g/m^3) due to a* 20% uncertainty *in individual measured variables and all measured variables combined. This is an instantaneous surface release measured at an elevation of* 1 foot. *There is an* inversion *present at* 150 feet.

	U	Q	σ_x	σ_y	σ_z	ALL
STAB: 4.5						
330 ft	0.679	0.034	0.021	0.035	0.035	0.709
1650 ft	0.017	4.74E-4	4.03E-4	9.57E-4	4.92E-4	0.018
3300 ft	9.66E-3	2.46E-5	1.96E-4	6.47E-5	2.55E-4	9.94E-3
8250 ft	4.35E-5	9.81E-8	1.62E-6	3.56E-8	1.64E-4	1.71E-4
STAB: 3.5						
330 ft	0.317	0.035	0.020	0.036	0.036	0.350
1650 ft	5.93E-3	3.43E-4	2.62E-4	1.90E-4	3.57E-4	6.23E-3
3300 ft	1.95E-3	9.85E-5	8.08E-5	5.97E-5	5.25E-5	2.03E-3
8250 ft	9.09E-5	9.75E-7	4.95E-6	6.83E-7	2.11E-4	2.33E-4
STAB: 1.5						
330 ft	0.022	3.29E-3	2.52E-3	3.42E-3	3.41E-3	0.027
1650 ft	3.85E-4	6.85E-5	5.42E-5	4.48E-5	1.46E-4	4.61E-4
3300 ft	1.40E-4	2.23E-5	1.81E-5	2.06E-5	8.60E-5	1.81E-4
8250 ft	1.71E-5	1.89E-6	2.31E-6	1.98E-6	4.04E-5	4.55E-5

Table 13: *Calculated concentration uncertainties (in g/m^3) due to a 20% uncertainty in individual measured variables and all measured variables combined. This is an instantaneous surface release measured at an elevation of 10 feet. There is an inversion present at 150 feet.*

	U	Q	σ_x	σ_y	σ_z	ALL
STAB: 4.5						
330 ft	0.594	0.029	0.018	0.031	0.022	0.620
1650 ft	0.017	4.59E-4	3.91E-4	9.27E-4	4.44E-4	0.017
3300 ft	9.54E-3	2.43E-4	1.94E-4	6.40E-5	2.44E-4	9.82E-3
8250 ft	4.35E-3	9.79E-6	1.61E-4	3.55E-6	1.66E-2	1.72E-2
STAB: 3.5						
330 ft	0.274	0.030	0.018	0.031	0.022	0.301
1650 ft	5.84E-3	3.38E-4	2.58E-4	1.87E-4	3.39E-4	6.13E-3
3300 ft	1.94E-3	9.80E-5	8.04E-5	5.94E-5	4.50E-5	2.02E-3
8250 ft	9.09E-5	9.75E-7	4.95E-6	6.83E-7	2.11E-4	2.33E-4
STAB: 1.5						
330 ft	0.021	3.14E-3	2.41E-3	0.011	2.94E-3	0.025
1650 ft			% UNCERTAINTIES ARE			
3300 ft			THE SAME			
8250 ft			AS THE 1FT LEVEL			

Table 14: *Calculated concentration uncertainties (in g/m^3) due to a 10% uncertainty in individual measured variables and all measured variables combined. This is a continuous release 2.5 m over water measured at an elevation of 1 m.*

	U	Q	H	σ_x	σ_y	σ_z	ALL
STAB: 4.5							
1113 m	6.57	0.15	4.8E-3	0.057	0.15	0.15	6.73
1000 m	0.035	2.3E-5	8.7E-7	4.6E-4	4.3E-5	2.2E-5	0.035
1750 m	2.15	0.029	4.8E-4	0.023	4.5E-3	0.029	2.17
2500 m	5.07E-5	2.9E-8	2.7E-10	1.3E-6	3.5E-8	2.9E-8	5.07E-5
STAB: 3.5							
1113 m	1.51	0.045	6.2E-4	0.035	0.046	0.045	1.53
1000 m	0.32	1.0E-3	1.8E-5	6.9E-3	3.7E-4	1.0E-3	0.32
1750 m	0.44	0.017	8.3E-5	9.1E-3	6.2E-3	0.012	0.44
2500 m	4.71E-3	1.3E-5	5.1E-8	1.6E-4	1.8E-6	1.3E-5	4.72E-3
STAB: 1.5							
1113 m	0.028	2.6E-5	1.9E-6	1.9E-3	2.6E-3	2.6E-3	0.029
1000 m	0.036	1.3E-3	1.5E-6	2.0E-3	8.6E-4	1.3E-3	0.038
1750 m	9.3E-3	7.9E-4	2.7E-7	6.1E-4	6.9E-4	7.9E-4	9.6E-3
2500 m	1.59E-3	5.6E-5	8.5E-9	1.4E-4	4.2E-5	5.6E-5	1.65E-3

Freeman *et al.* (1986) found uncertainties in the measured wind direction to be the largest contributor to the total measurement uncertainty of calculated concentrations for their Gaussian-plume simulations. Note that wind direction is not an explicit parameter in the Gaussian-plume equation. The downwind and crosswind distances, x and y, are expressed as functions of the wind direction vector for the Freeman *et al.* study. In that study, the receptor locations were specified such that they were aligned along the basic wind direction. The deviations from the basic wind direction represented by the uncertainty caused the plume centerline to be off the x-axis some distance, y', related to the angle of the deviation. The relationship of the downwind distance of the specified receptors to the transport wind would also be affected by the deviations in the wind direction. It is not clear how these effects were parameterized by Freeman *et al.* so that the partial derivatives in the propagation equation could be taken (the individual partial derivatives have been requested from the authors but are not currently available).

Freeman *et al.* (1986) interpret their results to say that uncertainty in the measured wind direction of 3 degrees contributes 40-50% of the total measurement uncertainty for downwind distances greater than 2 km at plume centerline. At these distances, the next largest contributors are sigma-y (20-30%) and sigma-z (10-20%). Wind speed is a relatively small contributor of uncertainty in this formulation (1-3%) and the source strength contribution to the uncertainty varied between 2% and 6%.

In Gaussian-puff formulae, such as that of AFTOX, the wind speed appears in the exponential term that represents the spread or transport of the puff along the x-axis. When the partial derivative with respect to u is taken, it becomes an important source of uncertainty in the propagation

equation (see Appendix A for the individual partial derivatives). Table 15 shows the calculated uncertainties from Table 7 expressed as percentages of the corresponding model-derived concentrations from Table 3. The size of the propagated uncertainty resulting from a 10% uncertainty in the wind speed relative to the propagated uncertainties for the other specified parameters is easily seen in this table. This table suggests that the wind speed uncertainty is much more critical to the dispersion calculation than uncertainties in the other parameters.

Freeman *et al.* used Monte Carlo statistical methods to determine if the propagation equation produced reasonable results for the specified uncertainties of their study. In general, they found the results reasonable for the ranges investigated with stability classes A and C. More stable classes tended to produce large inconsistances. Further investigation of the appropriateness of the propagation fromula for general applications was recommended by Freeman *et al.* (1986) and such an effort will be part of a future proposal by the present author to the Air Force Office of Scientific Research. In particular, since there are significant differences in results for the plume and puff formulations, the appropriateness of the uncertainty propagation equation for puff models, such as AFTOX, should be further considered to establish the validity of the present tentative results.

In Table 15, the results for the stability classes tested appear to be reasonably consistent, however, some inconsistencies are apparent at the greater distances. These may be the result of non-linearities in the Gaussian-puff equation or neglected terms in the propagation equation.

Pending further investigation, the uncertainty propagation method appears to be a useful way of quantifying the effects of measurement uncertainty. These results require further substantiation, perhaps using

Table 15: The calculated uncertainties shown in Table 7 expressed as percentages of the model-derived concentrations (Table 3).

	U	Q	σ_x	σ_y	σ_z	ALL
STAB: 4.5						
<100 m	128	10	6	10	10	132
1650 ft	105	5	4	10	5	108
3300 ft	61	3	2	0.5	3	63
STAB: 3.5						
330 ft	69	10	6	10	10	74
1650 ft	38	3	3	1	3	39
3300 ft	22	2	1	1	2	23
STAB: 1.5						
330 ft	31	5	3	5	5	34
1650 ft	0.7	0.2	0.1	0.1	0.2	0.8
3300 ft	4	1	10	1	1	5

Monte Carlo statistics as was done in Freeman *et al.* (1986) for the Gaussian-plume, but they suggest that puff models, like AFTOX, are highly sensitive to uncertainty in the specified wind speed and relatively insensitive to uncertainties in the other specified parameters of this study. While this result emphasizes the importance of using an accurately measured representative wind speed, it also suggests that uncertainties in the other specified parameters are much less important. Since these other parameters are usually not directly measureable, this would be a significant finding, because it would suggest that differences in the routines for determining the values of these parameters should not greatly influence the uncertainty of the model results.

Acknowledgements

Capt. Larry Key of AFESC and Dr. Bruce Kunkel of AFGL supplied the various forms of AFTOX used in this study and consultations throughout. Mr. D. L. Freeman of the Desert Research Institute supplied pre-publication copies of related papers.

The graduate student who did most of the computations in this project was Ms. Cathy E. Watson. The tables in this report are reproduced from her graduate research report.

Lt. Paul Beaudoin and Capt. Mike Moss, both of whom are AFIT students in the FSU Department of Meteorology, assisted Ms. Watson in the formulation and presentation of her report.

References

Freeman, D.L., R.T. Egami, N.F. Robinson and J.G. Watson, 1986: A method of propogating measurement uncertainties through dispersion models. JAPCA, 36, 246-253.

Hanna, S.A., G.A. Briggs, and R.P.Hosker, Jr., 1982: **Handbook on Atmospheric Diffusion**. Available from the National Technical Information Service as DE8200245 (DOE/TIC-11223).

Irwin, J.S., S.T. Rao, W.B. Petersen, and D.B. Turner, 1987: Relating error bounds for maximum concentration estimates to diffusion meteorology uncertainty. Paper submitted to **Atmospheric Environment**.

Pasquill, F., 1974: **Atmospheric Diffusion**, 2nd. ed., John Wiley & Sons, New York.

Turner, D.B., 1970: **Workbook of Atmospheric Dispersion Estimates**, National Air Pollution Control Administration, Cincinnati, Ohio.

APPENDIX A:

PARTIAL DERIVATIVES

The following is a list of the needed partial derivatives taken from the AFTOX Gaussian equation. For ease in derivation the following shorthand was used:

$$M = \frac{Q(t')}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z} \qquad N = e^{-1/2\left[\frac{x - U(t-t')}{\sigma_x}\right]^2}$$

$$O = e^{-1/2(y/\sigma_y)^2} \qquad P = \left[e^{-1/2\left[\frac{z-H}{\sigma_z}\right]^2} + e^{-1/2\left[\frac{z+H}{\sigma_z}\right]^2} \right]$$

$$A = e^{-\left[\frac{z-H}{\sigma_z}\right]^2} \qquad B = e^{-1/2\left[\frac{z+H}{\sigma_z}\right]^2}$$

where $G(x,y,z,t-t') = MNOP$.

The following simplifications were also used:

$$CZ = A \cdot \frac{(z-H)^2}{\sigma_z^3} + B \cdot \frac{(z+H)^2}{\sigma_z^3}, \qquad CT = x-U(t-t')$$

$$CZ1 = A \cdot \frac{(z-H)}{\sigma_z^2} - B \cdot \frac{(z+H)}{\sigma_z^2}, \qquad CT2 = \frac{(x-U(t-t'))(t-t')}{\sigma_x^2}$$

$$CT3 = \frac{(x-U(t-t'))^2}{\sigma_x^3}, \qquad YT1 = y^2/\sigma_y^2$$

$$\frac{\delta G}{\delta Q} = \frac{G}{Q}$$

$$\frac{\delta G}{\delta U} = G \cdot CT2$$

$$\frac{\delta G}{\delta \sigma_x} = \frac{-G}{\sigma_x} + (G \cdot CT3)$$

$$\frac{\delta G}{\delta \sigma_y} = \frac{-G}{\sigma_y} + (G \cdot \frac{YT1}{\sigma_y})$$

$$\frac{\delta G}{\delta \sigma_z} = \frac{-G}{\sigma_z} + (MNO \cdot CZ)$$

$$\frac{\delta G}{\delta H} = MNO \cdot CZ1$$

$$\frac{\delta^2 G}{\delta Q \delta U} = \frac{\delta G}{\delta Q} \cdot CT2$$

$$\frac{\delta^2 G}{\delta Q \delta \sigma_x} = \left[\frac{\delta G}{\delta Q} \cdot CT3 \right] - \left[\frac{\delta G / \delta Q}{\sigma_x} \right]$$

$$\frac{\delta^2 G}{\delta Q \delta \sigma_y} = \left[\frac{\delta G}{\delta Q} \cdot YT1 \right] - \left[\frac{\delta G / \delta Q}{\sigma_y} \right]$$

$$\frac{\delta^2 G}{\delta Q \delta \sigma_z} = (MNO/Q \cdot CZ) - \left[\frac{\delta G / \delta Q}{\sigma_z} \right]$$

$$\frac{\delta^2 G}{\delta Q \delta H} = MNO/Q \cdot CZ1$$

$$\frac{\delta^2 G}{\delta U \delta \sigma_x} = \left[\frac{-G}{\sigma_x} \cdot CT2 \right] + \left[G \cdot \left[(CT3 \cdot (t-t')) - (2 \cdot CT \cdot (t-t')/\sigma_x^3) \right] \right]$$

$$\frac{\delta^2 G}{\delta U \delta \sigma_y} = \frac{\delta G}{\delta \sigma_y} \cdot CT2$$

$$\frac{\delta^2 G}{\delta U \delta \sigma_z} = \frac{\delta G}{\delta \sigma_z} \cdot CT2$$

$$\frac{\delta^2 G}{\delta U \delta H} = \frac{\delta G}{\delta H} \cdot CT2$$

$$\frac{\delta^2 G}{\delta \sigma_x \delta \sigma_y} = \left(\frac{-G}{\sigma_x} \cdot YT1\right) + \frac{G}{\sigma_x \sigma_y} - \left(\frac{G}{\sigma_y} \cdot CT3\right) + (G \cdot CT3 \cdot YT1)$$

$$\frac{\delta^2 G}{\delta \sigma_x \delta \sigma_z} = \left(\frac{-G}{\sigma_z} \cdot CT3\right) + \frac{G}{\sigma_x \sigma_z} - (MNO/\sigma_x \cdot CZ) + (MNO \cdot CT3 \cdot CZ)$$

$$\frac{\delta^2 G}{\delta \sigma_x \delta H} = (-MNO/\sigma_x \cdot CZ1) + (MNO \cdot CT3 \cdot CZ1)$$

$$\frac{\delta^2 G}{\delta \sigma_y \delta \sigma_z} = \left(\frac{-G}{\sigma_z} \cdot YT1\right) + \frac{G}{\sigma_y \sigma_z} - (MNO/\sigma_y \cdot CZ) + (MNO \cdot YT1 \cdot CZ)$$

$$\frac{\delta^2 G}{\delta \sigma_y \delta H} = (-MNO/\sigma_y \cdot CZ1) + (MNO \cdot YT1 \cdot CZ1)$$

$$\frac{\delta^2 G}{\delta \sigma_z \delta H} = MNO \cdot \left[A \frac{(z-H)^3}{\sigma_z^5} - 2A \frac{(z-H)}{\sigma_z^3} - B \frac{(z+H)^3}{\sigma_z^5} + 2B \frac{(z+H)}{\sigma_z^3} \right] - (MNO/\sigma_z \cdot CZ1)$$

$$\frac{\delta^2 G}{\delta Q^2} = 0$$

$$\frac{\delta^2 G}{\delta U^2} = G \cdot \left[(CT2)^2 - \frac{(t-t')^2}{\sigma_x^2} \right]$$

$$\frac{\delta^2 G}{\delta \sigma_x^2} = G \cdot \left[\frac{2}{\sigma_x^2} - 2CT3/\sigma_x + (CT3)^2 - 3CT3/\sigma_x \right]$$

$$\frac{\delta^2 G}{\delta \sigma_y^2} = G \cdot \left[\frac{2}{\sigma_y^2} - 2YT1/\sigma_y + (YT1)^2 - 3YT1/\sigma_y \right]$$

$$\frac{\delta^2 G}{\delta \sigma_z^2} = \frac{2G}{\sigma_z^2} - \left(\frac{2MNO}{\sigma_z} \cdot CZ\right) + \left[MNO \cdot \left(A \left[\frac{(z-H)^4}{\sigma_z^6} - \frac{3(z-H)^2}{\sigma_z^4} \right] + B \left[\frac{(z+H)^4}{\sigma_z^6} - \frac{3(z+H)^2}{\sigma_z^4} \right] \right) \right]$$

$$\frac{\delta^2 G}{\delta H^2} = MNO \cdot \left(A \left[\frac{(z-H)^2}{\sigma_z^4} - \frac{1}{\sigma_z^2} \right] + B \left[\frac{(z+H)^2}{\sigma_z^4} - \frac{1}{\sigma_z^2} \right] \right)$$

INVERSION TERMS

The following are the necessary partials to be added to the P term whenever an inversion is present, including a shorthand version of frequently used expressions. The partials are summed along with the inversion terms and are added to their respective final uncertainty partials later in the program.

$CI = z-H-2NL \quad DI = z+H-2NL \quad EI = z-H+2NL \quad FI = z+H+2NL$

$$CI1 = \frac{(CI)^2}{\sigma_z^3} \qquad DI1 = \frac{(DI)^2}{\sigma_z^3}$$

$$EI1 = \frac{(EI)^2}{\sigma_z^3} \qquad FI1 = \frac{(FI)^2}{\sigma_z^3}$$

$$CI2 = CI/\sigma_z^2 \quad DI2 = DI/\sigma_z^2 \quad EI2 = EI/\sigma_z^2 \quad FI2 = FI/\sigma_z^2$$

$RMNO = PREEXP + XFAC + YFAC \qquad CIT2 = CIZ \cdot TRAVT/\sigma_x^2$

$CIZ = (x \cdot 1000 - U \cdot TRAVT) \qquad CIT3 = CIZ^2/\sigma_x^3$

$$\delta C/\delta\sigma_z = C \cdot CI1 \qquad \delta D/\delta\sigma_z = D \cdot DI1$$

$$\delta E/\delta\sigma_z = E \cdot EI1 \qquad \delta F/\delta\sigma_z = F \cdot FI1$$

$S1PZ = DCDSZ + DDDSZ + DEDSZ + DFDSZ$

$$\delta C/\delta H = C \cdot CI2 \qquad \delta D/\delta H = -D \cdot DI2$$

$$\delta E/\delta H = E \cdot EI2 \qquad \delta F/\delta H = -F \cdot FI2$$

$S1PH = DCDH + DDDH + DEDH + DFDH$

$$\delta^2 G/\delta Q \delta\sigma_z = (RMNO/MP) \cdot S1PZ$$

$$\delta^2 G/\delta Q \delta H = (RMNO/MP) \cdot S1PH$$

$$\delta^2 G/\delta U \delta\sigma_z = RMNO \cdot CIT2 \cdot S1PZ$$

$$\delta^2 G/\delta U \delta H = RMNO \cdot CIT2 \cdot S1PH$$

$$\delta^2 G/\delta\sigma_x \delta\sigma_z = (RMNO \cdot CIT3 \cdot S1PZ) - (RMNO/\sigma_x \cdot S1PZ)$$

$$\delta^2 G/\delta\sigma_x \delta H = (RMNO \cdot CIT3 \cdot S1PH) - (RMNO/\sigma_x \cdot S1PH)$$

$$\delta^2 G/\delta\sigma_y \delta\sigma_z = (RMNO \cdot Y^2/\sigma_y^3 \cdot S1PZ) - (RMNO/\sigma_y \cdot S1PZ)$$

$$\delta^2 G/\delta\sigma_y \delta H = (RMNO \cdot Y^2/\sigma_y^3 \cdot S1PH) - (RMNO/\sigma_y \cdot S1PH)$$

$$\delta^2 G/\delta\sigma_z \delta H = RMNO((C \cdot CI^3/\sigma_z^5 - 2C \cdot CI/\sigma_z^3 - D \cdot DI^3/\sigma_z^5 + 2D \cdot DI/\sigma_z^3 + E \cdot EI^3/\sigma_z^5 - 2E \cdot EI/\sigma_z^3 - F \cdot FI^3/\sigma_z^5 + 2F \cdot FI/\sigma_z^3) - (S1PH/\sigma_z)$$

$\delta^2C/\delta\sigma_z^2 = C \cdot (CI1^2 - 3CI1/\sigma_z)$ $\delta^2D/\delta\sigma_z^2 = D \cdot (DI1^2 - 3DI1/\sigma_z)$

$\delta^2E/\delta\sigma_z^2 = E \cdot (EI1^2 - 3EI1/\sigma_z)$ $\delta^2F/\delta\sigma_z^2 = F \cdot (FI1^2 - 3FI1/\sigma_z)$

S2PZ = D2CDSZ2 + D2DDSZ2 + D2EDSZ2 + D2FDSZ2

$\delta^2C/\delta H^2 = C \cdot (CI1/\sigma_z - 1/\sigma_z^2)$

$\delta^2D/\delta H^2 = D \cdot (DI1/\sigma_z - 1/\sigma_z^2)$

$\delta^2E/\delta H^2 = E \cdot (EI1/\sigma_z - 1/\sigma_z^2)$

$\delta^2F/\delta H^2 = F \cdot (FI1/\sigma_z - 1/\sigma_z^2)$

S2PH = D2CDH2 + D2DDH2 = D2EDH2 + D2FDH2

1987 AFOSR MINI GRANT RESEARCH PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by

Rodney C. Darrah
Program Director
MINI GRANT PROGRAM
Universal Energy Systems, Inc.
4401 Dayton-Xenia Road
Dayton, Ohio 45432
(513) 426-6900

FINAL REPORT

DESIGN OF H^{∞} MULTIVARIABLE OPTIMAL CONTROL SYSTEMS

Prepared by: Bor-Chin Chang

Academic Rank: Associate Professor

Department and University:

Before June 30, 1987: Department of Electrical and Computer Engineering and Technology
Bradley University, Peoria, IL 61625

After July 1, 1987: Department of Mechanical Engineering and Mechanics
Drexel University, Philadelphia, PA 19104
TEL: (215) 895-1790, 895-2352

Research Period: January 1987 -- December 1987

Date: January 19, 1988

Contract No: F49620-85-C-0013/SB5851-0360

DESIGN OF H^∞ MULTIVARIABLE OPTIMAL CONTROL SYSTEMS

by

Bor-Chin Chang

ABSTRACT

The objective of the research is to find efficient and reliable algorithms for H^∞ optimization problems. H^∞ optimization technique is a powerful tool which allows us to be able to handle the plant uncertainties and to consider a more general class of reference and disturbance inputs than that considered by Linear-Quadratic-Gaussian and Wiener-Hopf approaches. To make this powerful tool easy to use in the design of optimal control systems, the development of a complete computer-aided-design methodology based on this tool is necessary.

Several significant progresses have been made in this research period. These include "Size Reduction in Four-block H^∞ Formulation", "A Stable State-space Realization in the Formulation of H^∞ Norm Computation", "Fast Iterative Algorithm for Two-block H^∞ Optimization Problems", and "Optimal H^∞ Norm Computation for MIMO Systems with Multiple RHP Zeros".

In the paper "Size Reduction in Four-block H^∞ Formulation", some progress related to the size reduction of the H^∞ optimal controller will be presented. To reduce the order of the H^∞ optimal controller, the sizes of the state-space realizations of the rational matrices $R_{11}(s)$, $R_{12}(s)$, $R_{21}(s)$, and $R_{22}(s)$ in the four-block H^∞optimization problem formulation are required to be small. The recently discovered properties on the solution of the algebraic Riccati equation by Postlethwaite et. al. and the pole-zero cancellation technique by Doyle and Chu will be used to construct the minimal realizations of $R_{L1}(s)$, $R_{L2}(s)$, $R_{R1}(s)$, and $R_{R2}(s)$. From these realizations, the rational matrices $R_{11}(s)$, $R_{12}(s)$, $R_{21}(s)$, and $R_{22}(s)$ can be easily obtained. Only orthogonal transformations are involved in the computation, so the algorithm is numerically reliable.

Two problems are studied in the paper, "A Stable State-space Realization in the Formulation of H^∞ Norm Computation". In the two-block H^∞ optimization problem, usually we are given the state-space realizations of the proper rational matrices $R_1(s)$ and $R_2(s)$ whose poles are all in the open right half plane. The first is the evaluation of $\Phi(s)R_1(s)$ at $s = s_k$, k=1,2,...,n, where $\Phi(s)$ is an inner function whose zeros (s_k, k=1,2,...,n) are the poles of $R_1(s)$. This evaluation is essential if Chang

and Pearson's method is used for computing the optimal H^∞ norm. The problem is solved in state space via the solutions of Lyapunov equations. Neither polynomial matrix manipulations nor numerical pole-zero cancellations are involved in the evaluation. The second problem is to find a stable state-space realization of $S(s) = U(s)R_2(s)$ where U(s) is an inner matrix. This problem arises in the spectral factorization of $\gamma^2 I - R_2^* R_2$. Doyle and Chu had a method for constructing stable S(s) based on a minimal realization of $R_2(s)$. An alternate method is proposed. The alternate method does not require a minimal realization of $R_2(s)$ and only a Lyapunov equation is involved.

In the paper "Fast Iterative Algorithm for Two-block H^∞ Optimization Problems", we developed a fast iterative algorithm for computing the optimal norm of the two-block H^∞ optimization problem. The two-block H^∞ optimization problem arises in the optimal disturbance attenuation with control weighting, the minimization of a weighted sum involving both the sensitivity function and its complement, and any control problems with more controlled outputs than control inputs. In this optimization problem, the most computationally demanding work is the computation of the optimal H^∞ norm denoted by γ_o. The problem of computing γ_o can be considered as that of finding a γ such that $\mu(\gamma)$ equals to 1. Some new properties of $\mu(\gamma)$ are revealed and studied. Based on these properties and those found by Chu and Doyle, and Wang and Pearson, a very fast search scheme for finding γ_o is proposed. This algorithm requires less computation and converges much more quickly than any other existing method.

In either two-block or four-block H^∞ optimization problems, we need to compute the optimal H^∞ norm of the regular (one-block) H^∞ optimization problem in each iteration. For this computation, Chang and Pearson's approach is used because of its simplicity. The only drawback of the Chang and Pearson's method is the restriction that the RHP zeros should be distinct. This restriction was removed in the paper "Optimal H^∞ Norm Computation for MIMO Systems with Multiple RHP Zeros". The MIMO case of the one-block optimal H^∞-norm computation with multiple RHP zeros can be also reduced to a simple eigenvalue or singular value computation problem.

ACKNOWLEDGMENTS

I would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsoring my research. I also would like to express my appreciation to Dr. Siva Banda, Lt. Tim McQuade, and Dr. Hsi-Han Yeh of AFWAL/FIGCA, Wright-Patterson AFB for their continuous encouragement and providing valuable comments. Thanks are also due to the personnels in Universal Energy Systems, Inc. for their assistance in conducting the Mini Grant Research Program.

I. INTRODUCTION

In the reserach period, several significant progresses have been made. These progresses are summarized in the following papers: "Size Reduction in Four-block H^∞ Formulation" [1], "A Stable State-space Realization in the Formulation of H^∞ Norm Computation" [2], "Fast Iterative Algorithm for Two-block H^∞ Optimization Problems" [3], and "Optimal H^∞ Norm Computation for MIMO Systems with Multiple RHP Zeros" [4].

The major contributions of these progresses will be presented in Sections III-VI. In the rest of the section we will briefly review the concept of H^∞ optimization and its potential to be a powerful tool in control systems design. Section II shows how a control problem can be formulated as an H^∞ optimization problem. Conclusions and recommendations for further research are given in Section VII.

The concept of H^∞ optimization was initiated by G. Zames [5] in 1981. Since then, this new research area has attracted many researchers [6-23]. The popularity of this research area is mainly because this area appears to have a major impact on future engineering practice.

Two robustness issues are important in control system design. One is the command/ disturbance input uncertainties and the other is the plant uncertainties. The controller is to be designed such that the closed-loop system remains stable for all possible plant perturbations and the error response is admissible for every disturbance and command in the prescribed set under all possible plant perturbations.

The prescribed set of disturbances and commands is modeled by the designer according to the actual environment of the system. If the set of disturbances and commands under consideration can be described by a random process then a stochastic design approach like LQG [24] or Wiener-Hopf [25] method can be used to minimize the mean value of the error energy. On the other hand, if we do not have any useful statistical information about the disturbances and the commands then a new developed H^∞ optimization approach is recommended. The H^∞ optimization approach is a minmax design method which minimizes the maximal error energy for all possible disturbances and commands in the prescribed set.

H^∞ optimization technique is a powerful tool which not only allows us to consider a more general class of command and disturbance inputs than those considered by LQG and W-H

approaches but also has the ability to handle the plant uncertainties. H^∞ optimization technique uses the H^∞ norm of a transfer function matrix as the measure of the error response of a system. This reflects the maximum error energy which may occur in reality. YJB controller parametrization [25-27,13,14] and regulator design techniques [28-30] are employed to characterize the set of the controllers which solve the stabilization and regulator problems. Then among the controllers in the set, we will find the one that minimizes the error response subject to control-input and stability-margin constraints.

The control problem can be formulated in a very general way by using H^∞ optimization concept. The models of the plant uncertainties and the reference and disturbance inputs are general enough to reflect the situations found in actual problems. Also, the controller structure is the most general under linear time-invariance. The two-degree-of-freedom controllers are included in this structure.

Although the H^∞ design basically is a frequency-domain approach, the computation is not necessarily done in frequency domain. As a matter of fact, many algorithms developed on the state-space framework have better numerical properties than their counterparts based on polynomial matrix manipulations. All the computations in our design are implemented on the state-space framework. Furthermore, many pole-zero cancellations can be done theoretically without using any numerical minimal realization technique. This theoretical pole-zero cancellation technique [13,14,22,23,1] can save a lot of computing time and avoid incomplete cancellation due to rounding errors. It is possible to perform all the computation in a numerically stable, reliable, and efficient way.

II. H^∞ FORMULATION OF CONTROL PROBLEMS

Consider the system

$$y(s) = P(s)u(s) + d(s) \tag{2-1a}$$

$$u(s) = - Q(s)y(s) \tag{2-1b}$$

The nominal plant transfer function P(s) is given and the set of disturbances is assumed to be any signal vector with energy less than or equal to one. The objective is to find a controller Q(s) such that the closed-loop system is stable and the maximal energy of the disturbance response is minimized subject to stability-margin and control input constraints.

The transfer function from d(s) to y(s) is given by

$$[\, I + P(s)\, Q(s)\,]^{-1} \tag{2-2}$$

which is the sensitivity function of the closed-loop system. From [31], we have

$$\| [I+PQ]^{-1} \|_\infty := \sup_\omega \bar{\sigma} [\, (I+PQ)^{-1}(j\omega)\,] = \sup \{\, \|y\|_2 \mid \|d\|_2 \le 1 \,\} \tag{2-3}$$

where

$$\| \zeta \|_2^2 = \int_0^\infty \zeta(t)^T \zeta(t)\, dt \tag{2-4}$$

Therefore, minimizing $\|\Phi\|_\infty$ is equivalent to minimizing the maximum $\|y\|_2$ for all d with $\|d\|_2 \le 1$. That is, the maximal energy of the disturbance response is equivalent to the H^∞ norm of $[I+PQ]^{-1}$. Define

$$\Phi = \begin{bmatrix} W_1 (I+PQ)^{-1} \\ W_2 PQ (I+PQ)^{-1} \end{bmatrix} \tag{2-5}$$

where $W_1(I+PQ)^{-1}$ is the weighted sensitivity function and $W_2PQ(I+PQ)^{-1}$ is the weighted complementary sensitivity function. W_1 and W_2 are weighting matrices chosen by the designer according to the actual environment.

The problem of finding a Q(s) which stabilizes the closed-loop system and minimizes $\|\Phi\|_\infty$ can be reformulated as follows. Consider the following system

$$\begin{bmatrix} z_1 \\ z_2 \\ y \end{bmatrix} = \begin{bmatrix} W_1 & W_1P \\ 0 & -W_2P \\ I & P \end{bmatrix} \begin{bmatrix} d \\ u \end{bmatrix} \tag{2-6a}$$

$$u = -Qy \tag{2-6b}$$

It is easy to show that the matrix Φ defined by (2-5) is just the transfer function from d to $[z_1^T \;\; z_2^T]^T$ of the closed-loop system (2-6).

The system (2-6) is just a special case of the following

$$\begin{bmatrix} z(s) \\ y(s) \end{bmatrix} = \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix} \begin{bmatrix} v(s) \\ u(s) \end{bmatrix} := G(s) \begin{bmatrix} v(s) \\ u(s) \end{bmatrix} \tag{2-7a}$$

$$u(s) = Q(s)y(s) \tag{2-7b}$$

where $G_{11}(s) \in \mathbb{R}(s)^{pxq}$, $G_{12}(s) \in \mathbb{R}(s)^{pxm}$, $G_{21}(s) \in \mathbb{R}(s)^{rxq}$, and $G_{22}(s) \in \mathbb{R}(s)^{rxm}$. The H^∞ optimization problem is to find a proper controller Q(s) such that the closed-loop system is internally stable and $\|\Phi\|_\infty$ is minimized where

$$\Phi(s) = G_{11}(s) + G_{12}(s)\, Q(s)\, [\, I - G_{22}(s)\, Q(s)\,]^{-1}\, G_{21}(s) \tag{2-8}$$

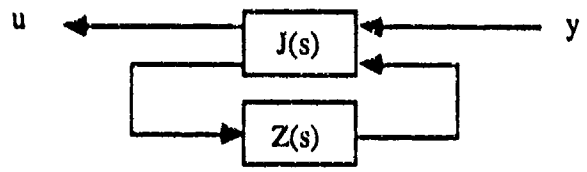
That is, $\Phi(s)$ is the transfer function of the closed-loop system from v to z. Let

$$G(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right] \tag{2-9}$$

be a minimal realization of G(s). Here

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \{ A, B, C, D \} \tag{2-10}$$

implies a state-space realization and $G(s) = D + C\,(sI - A)^{-1}\,B$. Suppose the realization $\{ A, B_2, C_2, D_{22} \}$ is stabilizable and detectable, then by using Doyle's version [13,14] of YJB controller parametrization [25,26], the set of stabilizing controllers can be characterized by the following structure



with $Z(s) \in (RH^\infty)^{mxr}$ and $I+D_{22}Z(\infty)$ invertible.

Fig. 2-1 Parametrization of stabilizing controllers.

where

$$J(s) = \left[\begin{array}{c|cc} A+B_2F+KC_2+KD_{22}F & -K & B_2+KD_{22} \\ \hline F & 0 & I \\ -(C_2 + D_{22}F) & I & -D_{22} \end{array}\right] \tag{2-11}$$

and F and K are chosen such that eigenvalues of $A+B_2F$ and $A+KC_2$ are stable and Z(s) can be any proper stable rational. By using this parametrization, $\Phi(s)$ can be rewritten as [13,14]

$$\Phi(s) = H(s) + S(s)Z(s)T(s) \tag{2-12a}$$

where

$$H(s) = \left[\begin{array}{cc|c} A+B_2F & -B_2F & B_1 \\ 0 & A+KC_2 & B_1+KD_{21} \\ \hline C_1+D_{12}F & -D_{12}F & D_{11} \end{array}\right] \tag{2-12b}$$

$$S(s) = \left[\begin{array}{c|c} A+B_2F & B_2 \\ \hline C_1+D_{12}F & D_{12} \end{array}\right] \tag{2-12c}$$

$$T(s) = \left[\begin{array}{c|c} A+KC_2 & B_1+KD_{21} \\ \hline C_2 & D_{21} \end{array}\right] \tag{2-12d}$$

If S and T are thin and fat respectively, i.e., $p > m$ and $r < q$ and they can be decomposed as

$$S(s) = S_i(s)S_o(s), \qquad T(s) = T_o(s)T_i(s) \tag{2-13}$$

where $S_i \in (RH^\infty)^{p\times m}$ is inner, $T_i \in (RH^\infty)^{r\times q}$ is co-inner, $S_o \in (RH^\infty)^{m\times m}$ and $T_o \in (RH^\infty)^{r\times r}$ are outer matrices. Furthermore, we can find a $S_\perp(s) \in (RH^\infty)^{p\times(p-m)}$ such that $[\,S_i \;\; S_\perp\,]$ is square and inner. Similarly, $T_\perp(s) \in (RH^\infty)^{(q-r)\times q}$ can be found such that $[\,T_i^T \;\; T_\perp^T\,]^T$ is square and co-inner.

Now, (2-12a) can be rewritten as

$$\Phi = H + \begin{bmatrix} S_i & S_\perp \end{bmatrix} \begin{bmatrix} \hat{Z} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} T_i \\ T_\perp \end{bmatrix} \tag{2-14}$$

here $\hat{Z} = S_o Z T_o$. Define $\hat{\Phi} = [\,S_i \;\; S_\perp\,]^* \Phi [\,T_i^* \;\; T_\perp^*\,]$, then we have

$$\hat{\Phi} = \begin{bmatrix} S_i^*HT_i^* + \hat{Z} & S_i^*HT_\perp^* \\ S_\perp^*HT_i^* & S_\perp^*HT_\perp^* \end{bmatrix} := \begin{bmatrix} R_{11} + \hat{Z} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \tag{2-15}$$

Note that $\| \hat{\Phi} \|_\infty = \| \Phi \|_\infty$ and therefore the problem becomes to:

Find a $\hat{Z}(s) \in (RH^\infty)^{mxr}$ such that $\| \hat{\Phi} \|_\infty$ is minimized
where $\hat{\Phi}$ is given by (2-15). (2-16)

Problem (2-16) is referred to as the four-block H^∞ optimization problem.

The two-block H^∞ optimization problem is just a special case of problem (2-16) and the corresponding $\hat{\Phi}(s)$ is a special case of (2-15) with either $[R_{21} \; R_{22}] = 0$ or $[R_{12}^T \; R_{22}^T] = 0$.

The one-block H^∞ optimization problem was solved by Francis, Helton, and Zames [7], Chang and Pearson [8], Safonov and Verma [9], and Glover [10]. An iterative type algorithm was originally proposed by Doyle [12] for solving the four-block and two-block H^∞ optimization problems. The major burden of the computation is in the evaluation of the optimal H^∞ norm. Chang, Banda, and McQuade [3] proposed a fast iterative algorithm for computing the optimal H^∞ norm of the two-block problem. For most problems we encountered, the optimal two-block H^∞ norm can be accurately (to a double precision accuracy) obtained within three or four iterations. The computation of the four-block H^∞ norm is much more complicated than two-block problems. Chang and Pearson [19] and Chu, Doyle and Lee [20] greatly simplified the original Doyle's four-block H^∞ norm computation algorithm.

III. SIZE REDUCTION IN FOUR-BLOCK H^∞ FORMULATION

In the paper "Size Reduction in Four-Block H^∞ Formulation" [1], some progress related to the size reduction of the H^∞ optimal controller is presented. The four-block problem arises when the exogenous inputs are more than the measured outputs and the controlled outputs are more than the control inputs. The central part of the four-block H^∞ optimization problem is to solve the following,

$$\inf_{Z(s) \in H^\infty} \left\| \begin{bmatrix} R_{11}(s) + Z(s) & R_{12}(s) \\ R_{21}(s) & R_{22}(s) \end{bmatrix} \right\|_\infty \tag{3-1}$$

where $R_{11}(s)$, $R_{12}(s)$, $R_{21}(s)$ and $R_{22}(s)$ are completely unstable proper rational matrices determined by the given plant through Youla's controller parametrization [25,26,13,14], inner-outer factorizations [13,14], and some matrix manipulations. $Z(s)$ is a stabilizing parameter matrix to be

found such that the above H^∞ norm is minimized. To reduce the complexity of the H^∞ optimal controller, usually we would like the size of the state-space realization of Z(s) to be small. It is well known that smaller sizes of $R_{11}(s)$, $R_{12}(s)$, $R_{21}(s)$ and $R_{22}(s)$ are of help to decrease the order of Z(s).

$R_{11}(s)$, $R_{12}(s)$, $R_{21}(s)$ and $R_{22}(s)$ are constructed by cascading several rational matrices. If no pole-zero cancellation is performed during the construction, the orders of $R_{11}(s)$, $R_{12}(s)$, $R_{21}(s)$ and $R_{22}(s)$ usually are artificially high. The size reduction of $R_{11}(s)$, $R_{12}(s)$, $R_{21}(s)$ and $R_{22}(s)$ has been considered by Doyle and Chu [13,14], Limebeer and Hung [22], and Postlethwaite, Gu, and O'Young [23]. Doyle and Chu used state-space similarity transformation techniques to cancel some poles and zeros and derived state-space realizations for $R_L(s)$ and $R_R(s)$ such that

$$R(s) = \begin{bmatrix} R_{11}(s) & R_{12}(s) \\ R_{21}(s) & R_{22}(s) \end{bmatrix} = R_L(s)\, R_R(s) \tag{3-2}$$

The state-space realizations of $R_L(s)$ and $R_R(s)$ are given by

$$R_L(s) = \left[\begin{array}{c|cc} -(A+B_2F)^T & (C_1+D_{12}F)^T & -XK \\ \hline -(B_2R_D^{-1/2})^T & (D_{12}R_D^{-1/2})^T & 0 \\ D_\perp^T C_1 X^\dagger & D_\perp^T & 0 \end{array}\right] \tag{3-3a}$$

$$R_R(s) = \left[\begin{array}{c|cc} -(A+KC_2)^T & -(\tilde{R}_D^{-1/2}C_2)^T & Y^\dagger B_1 \tilde{D}_\perp^T \\ \hline C_1Y+D_{11}(B_1+KD_{21})^T & D_{11}(\tilde{R}_D^{-1/2}D_{21})^T & D_{11}\tilde{D}_\perp^T \\ 0 & \tilde{R}_D^{1/2} & 0 \end{array}\right] \tag{3-3b}$$

where

$$F = -R_D^{-1}(D_{12}^T C_1 + B_2^T X) \tag{3-4a}$$

$$R_D = D_{12}^T D_{12} \tag{3-4b}$$

$$X = \mathrm{Ric}\begin{bmatrix} A - B_2R_D^{-1}D_{12}^TC_1 & -B_2R_D^{-1}B_2^T \\ -C_1^TD_\perp D_\perp^TC_1 & -(A - B_2R_D^{-1}D_{12}^TC_1)^T \end{bmatrix} \tag{3-4c}$$

$D_\perp$ is the orthogonal complement of $D_{12}R_D^{-1/2}$ so that

$$\begin{bmatrix} D_{12}R_D^{-1/2} & D_\perp \end{bmatrix} \text{ is square and orthogonal.} \tag{3-4d}$$

$$(R_D^{1/2})^TR_D^{1/2} = R_D \tag{3-4e}$$

and

$$K = -(B_1D_{21}^T + YC_2^T)\tilde{R}_D^{-1} \tag{3-5a}$$

$$\tilde{R}_D = D_{21}D_{21}^T \tag{3-5b}$$

$$Y = \mathrm{Ric}\begin{bmatrix} (A - B_1D_{21}^T\tilde{R}_D^TC_2)^T & -C_2^T\tilde{R}_D^{-1}C_2 \\ -B_1\tilde{D}_\perp^T\tilde{D}_\perp B_1^T & -(A - B_1\tilde{D}_\perp^T\tilde{R}_D^{-1}C_2) \end{bmatrix} \tag{3-5c}$$

$\tilde{D}_\perp$ is the orthogonal complement of $\tilde{R}_D^{-1/2}D_{21}$ so that

$$\begin{bmatrix} \tilde{R}_D^{-1/2}D_{21} \\ \tilde{D}_\perp \end{bmatrix} \text{ is square and orthogonal.} \tag{3-5d}$$

$$(\tilde{R}_D^{1/2})(\tilde{R}_D^{1/2})^T = \tilde{R}_D \tag{3-5e}$$

Neither $R_L(s)$ nor $R_R(s)$ is a minimal realization. Limebeer and Hung derived a minimal realization of $R_{11}(s)$ in the one-block case (the case with $R_{12}(s)$, $R_{21}(s)$ and $R_{22}(s)$ being zero matrices) for finding a bound on the McMillan degree of the H^∞ optimal controller. This minimal realization process is not numerically reliable although it is adequate for their theoretical purpose. Recently, Postlethwaite, Gu, and O'Young unveiled some useful properties on the solution of the algebraic Riccati equation and used them to develop a numerically reliable algorithm for the minimal realization of $R_{11}(s)$ in the one-block case.

In [1] we consider the most general case: the four-block case, i.e., the case with $[R_{21}(s)\ R_{22}(s)] \neq 0$ and $[R_{12}^T(s)\ R_{22}^T(s)] \neq 0$. One-block and two-block (the case with $[R_{21}(s)\ R_{22}(s)] = 0$ or $[R_{12}^T(s)\ R_{22}^T(s)] = 0$) problems are just special cases of the four-block problem. We start from

the state-space realizations of $R_L(s)$ and $R_R(s)$ which were derived by Doyle and Chu. Partitioning $R_L(s)$ and $R_R(s)$ as

$$R_L(s) = \begin{bmatrix} R_{L1}(s) \\ R_{L2}(s) \end{bmatrix}, \qquad R_R(s) = \begin{bmatrix} R_{R1}(s) & R_{R2}(s) \end{bmatrix} \tag{3-6a}$$

then we have

$$\begin{aligned} R_{11}(s) &= R_{L1}(s)\, R_{R1}(s), \quad R_{12}(s) = R_{L1}(s)\, R_{R2}(s) \\ R_{21}(s) &= R_{L2}(s)\, R_{R1}(s), \quad R_{22}(s) = R_{L2}(s)\, R_{R2}(s) \end{aligned} \tag{3-6b}$$

where

$$R_{L1}(s) = \left[\begin{array}{c|cc} -(A+B_2F)^T & (C_1+D_{12}F)^T & -XK \\ \hline -(B_2R_D^{-1/2})^T & (D_{12}R_D^{-1/2})^T & 0 \end{array}\right] \tag{3-6c}$$

$$R_{L2}(s) = \left[\begin{array}{c|cc} -(A+B_2F)^T & (C_1+D_{12}F)^T & -XK \\ \hline D_\perp^T C_1 X^\dagger & D_\perp^T & 0 \end{array}\right] \tag{3-6d}$$

$$R_{R1}(s) = \left[\begin{array}{c|c} -(A+KC_2)^T & -(\tilde{R}_D^{-1/2} C_2)^T \\ \hline C_1Y+D_{11}(B_1+KD_{21})^T & D_{11}(\tilde{R}_D^{-1/2} D_{21})^T \\ 0 & \tilde{R}_D^{1/2} \end{array}\right] \tag{3-6e}$$

$$R_{R2}(s) = \left[\begin{array}{c|c} -(A+KC_2)^T & Y^\dagger B_1 \tilde{D}_\perp^T \\ \hline C_1Y+D_{11}(B_1+KD_{21})^T & D_{11}\tilde{D}_\perp^T \\ 0 & 0 \end{array}\right] \tag{3-6f}$$

The recently discovered properties on the solution of the algebraic Riccati equation by Postlethwaite et. al. [23] and the pole-zero cancellation technique by the similarity transformation are used for constructing the minimal realizations of $R_{L1}(s)$, $R_{L2}(s)$, $R_{R1}(s)$, and $R_{R2}(s)$. Only orthogonal

transformations are involved in the computation, so the algorithm is numerically reliable. The new result of Postlethwaite et. al. is listed as follows.

Theorem 3.1: [23]

Consider the ARE,

$$A^TX + XA - XBB^TX + C^TC = 0 \tag{3-7a}$$

There exists an orthogonal matrix $U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}$ such that

$$U^TAU = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix} \tag{3-7b}$$

$$U^TB = \begin{bmatrix} U_1^TB \\ U_2^TB \end{bmatrix}, \qquad CU = \begin{bmatrix} 0 & CU_2 \end{bmatrix} \tag{3-7c}$$

and

$$U^TXU = \begin{bmatrix} 0 & 0 \\ 0 & X_2 \end{bmatrix} \tag{3-7d}$$

The eigenvalues of A_1 are unobservable but stable. That is, the set of eigenvalues of A_3 includes all the observable and all the unstable eigenvalues of A. The rank of X_2, and therefore the rank of X, equals the dimension of A_3. $X_2 > 0$ satisfies the following ARE,

$$A_3^TX_2 + X_2A_3 - X_2U_2^TBB^TU_2X_2 + U_2^TC^TCU_2 = 0 \tag{3-7e}$$

The algorithm we developed is different from that of Postlethwaite et. al.. The advantage of the decomposition $R(s) = R_L(s)R_R(s)$ is taken to simplify the size reduction process for $R_{11}(s)$, $R_{12}(s)$, $R_{21}(s)$ and $R_{22}(s)$. Although we do not try to construct minimal realizations for $R_{ij}(s)$, we do have minimal realizations for $R_{Li}(s)$ and $R_{Rj}(s)$, i,j=1,2, and there is no possible mathematically identical pole-zero cancellation between $R_{Li}(s)$ and $R_{Rj}(s)$.

In the following, we will derive minimal state-space realizations for $R_{L1}(s)$ and $R_{L2}(s)$. Minimal realizations of $R_{R1}(s)$ and $R_{R2}(s)$ can be obtained in a similar way. First of all, let us consider $R_{L1}(s)$.

Minimal Realization of $R_{L1}(s)$

The first step is to apply Theorem 3.1 to the solution of the algebraic Riccati equation (3-4c). There exists an orthogonal matrix

$$U = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \tag{3-8a}$$

such that

$$U^T(A - B_2R_D^{-1}D_{12}^TC_1)U = \begin{bmatrix} A_1 & A_{12} \\ 0 & A_2 \end{bmatrix} \tag{3-8b}$$

$$U^TB_2R_D^{-1/2} = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix}, \quad D_\perp^TC_1U = \begin{bmatrix} 0 & W_2 \end{bmatrix} \tag{3-8c}$$

and

$$U^TXU = \begin{bmatrix} 0 & 0 \\ 0 & X_2 \end{bmatrix} \tag{3-8d}$$

where X_2 is the positive definite solution of the following reduced-order algebraic Riccati equation

$$A_2^TX_2 + X_2A_2 - X_2L_2L_2^TX_2 + W_2^TW_2 = 0 \tag{3-8e}$$

Applying the similarity transformation (U^T, U) to the realization of $R_{L1}(s)$ in (3-6c) and then using (3-4) and (3-8), we have

$$R_{L1}(s) = \left[\begin{array}{cc|cc} -A_1^T & 0 & 0 & 0 \\ -(A_{12}-L_1L_2^TX_2)^T & -(A_2-L_2L_2^TX_2)^T & (D_\perp W_2 - D_{12}R_D^{-1}B_2^TU_2X_2)^T & -X_2U_2^TK \\ \hline -L_1^T & -L_2^T & (D_{12}R_D^{-1/2})^T & 0 \end{array}\right]$$

$$= \left[\begin{array}{cc|c} -(A_2-L_2L_2^TX_2)^T & (D_\perp W_2 - D_{12}R_D^{-1}B_2^TU_2X_2)^T & -X_2U_2^TK \\ \hline -L_2^T & (D_{12}R_D^{-1/2})^T & 0 \end{array}\right] \tag{3-9}$$

The size of the state-space realization of $R_{L1}(s)$ has been reduced from the dimension of A to that of A_2 and the realization in (3-9) is controllable according to the following.

Remark 3.1:

The state-space realization of $R_{L1}(s)$ in (3-9) is controllable.

Proof: See [1].

Remark 3.2:

If $R_{L2}(s) = 0$, i.e., $D_\perp = 0$, then it is easy to prove that the pair $\{L_2^T, (A_2 - L_2L_2^TX_2)^T\}$ is observable. In this case, the realization of $R_{L1}(s)$ in (3-9) is already a minimal realization.

Now the realization of $R_{L1}(s)$ in (3-9) is controllable but not necessarily observable if $R_{L2}(s) \neq 0$. By using the Van Dooren's version [32] of Rosenbrock's algorithm [33], we can find an orthogonal matrix

$$V = \begin{bmatrix} V_1 & V_2 \end{bmatrix} \tag{3-10a}$$

such that

$$\begin{aligned} & V^T(A_2 - L_2L_2^TX_2)^T V \\ &= \begin{bmatrix} -V_1^T(A_2 - L_2L_2^TX_2)^TV_1 & 0 \\ -V_2^T(A_2 - L_2L_2^TX_2)^TV_1 & -V_2^T(A_2 - L_2L_2^TX_2)^TV_2 \end{bmatrix} \end{aligned} \tag{3-10b}$$

and

$$-L_2^TV = \begin{bmatrix} -L_2^TV_1 & 0 \end{bmatrix} \tag{3-10c}$$

and the pair

$$(-L_2^TV_1,\ -V_1^T(A_2 - L_2L_2^TX_2)^TV_1) \tag{3-10d}$$

is observable.

Applying the similarity transformation (V^T, V) to the realization of $R_{L1}(s)$ in (3-9) and eliminating the unobservable part, we have the following minimal realization

$$R_{L1}(s) = \left[\begin{array}{c|cc} -V_1^T(A_2 - L_2L_2^TX_2)^TV_1 & V_1^T(D_\perp W_2 - D_{12}R_D^{-1}B_2^TU_2X_2)^T & -V_1^TX_2U_2^TK \\ \hline -L_2^TV_1 & (D_{12}R_D^{-1/2})^T & 0 \end{array}\right] \tag{3-11}$$

Minimal Realization of $R_{L2}(s)$

The first step is just similar to that of the minimal realization of $R_{L1}(s)$. Applying the similarity transformation (U^T, U) to the realization of $R_{L2}(s)$ in (3-6d), and then using (3-4) and (3-8), we have

$$R_{L2}(s) = \left[\begin{array}{cc|cc} -A_1^T & 0 & 0 & 0 \\ -(A_{12}-L_1L_2^TX_2)^T & -(A_2-L_2L_2^TX_2)^T & (D_\perp W_2 - D_{12}R_D^{-1}B_2^TU_2X_2)^T & -X_2U_2^TK \\ \hline 0 & W_2X_2^{-1} & D_\perp^T & 0 \end{array}\right]$$

$$= \left[\begin{array}{c|cc} -(A_2-L_2L_2^TX_2)^T & (D_\perp W_2 - D_{12}R_D^{-1}B_2^TU_2X_2)^T & -X_2U_2^TK \\ \hline W_2X_2^{-1} & D_\perp^T & 0 \end{array}\right] \qquad (3\text{-}12)$$

For the same reason of Remark 3.1, the realization of $R_{L2}(s)$ in (3-12) is controllable. Next, we can find an orthogonal matrix

$$S = \begin{bmatrix} S_1 & S_2 \end{bmatrix} \qquad (3\text{-}13a)$$

such that

$$\begin{aligned} & S^T(A_2 - L_2L_2^TX_2)^T S \\ &= \begin{bmatrix} S_1^T(A_2 - L_2L_2^TX_2)^TS_1 & 0 \\ S_2^T(A_2 - L_2L_2^TX_2)^TS_1 & S_2^T(A_2 - L_2L_2^TX_2)^TS_2 \end{bmatrix} \end{aligned} \qquad (3\text{-}13b)$$

and

$$W_2X_2^{-1}S = \begin{bmatrix} W_2X_2^{-1}S_1 & 0 \end{bmatrix} \qquad (3\text{-}13c)$$

and the pair

$$\{\, W_2X_2^{-1}S_1,\ S_1^T(A_2 - L_2L_2^TX_2)^TS_1 \,\} \qquad (3\text{-}13d)$$

is observable.

Applying the similarity transformation (S^T, S) to the realization of $R_{L2}(s)$ in (3-12) and eliminating the unobservable part, we have the following minimal realization

$$R_{L2}(s) = \left[\begin{array}{c|cc} -S_1^T(A_2-L_2L_2^TX_2)^TS_1 & S_1^T(D_\perp W_2 - D_{12}R_D^{-1}B_2^TU_2X_2)^T & -S_1^TX_2U_2^TK \\ \hline W_2X_2^{-1}S_1 & D_\perp^T & 0 \end{array}\right] \qquad (3\text{-}14)$$

IV. A STABLE STATE-SPACE REALIZATION IN THE FORMULATION OF H^∞ NORM COMPUTATION

In the two-block H^∞ optimization problem, usually we are given the state-space realizations of the proper rational matrices $R_1(s)$ and $R_2(s)$ whose poles are all in the open right half plane. That is, the two-block H^∞ optimization problem includes the computation of

$$\inf_{K(s) \in (RH^\infty)^{mxr}} \left\| \begin{bmatrix} R_1(s) + K(s) \\ R_2(s) \end{bmatrix} \right\|_\infty := \gamma_0 \qquad (4\text{-}1)$$

and the construction of an optimal $K(s)$. $R_1(s)$ and $R_2(s)$ are in $(RL^\infty)^{mxr}$ and $(RL^\infty)^{qxr}$ respectively. The $R_1(s)$, $R_2(s)$ and $K(s)$ in (4-1) are just the same as the $R_{11}(s)$, $R_{21}(s)$ and $Z(s)$ in (3-1) with $[R_{12}(s)^T \; R_{22}(s)^T] = 0$.

Let $\phi(s)$ be an inner function whose zeros $\{s_k, k=1,2,...,n\}$ are the poles of $R_1(s)$ and $U(s)$ be an inner matrix such that $N(s) = \phi(s)R_1(s)$ and $S(s) = U(s)R_2(s)$ are stable. Then it is easy to see that the problem of (4-1) is the same as that of the following

$$\inf_{K(s) \in (RH^\infty)^{mxr}} \left\| \begin{bmatrix} N(s) + \phi(s) K(s) \\ S(s) \end{bmatrix} \right\|_\infty := \gamma_0 \qquad (4\text{-}2)$$

In solving the two-block H^∞ optimization problem, we need to compute the H^∞ norm of the one-block problem for several times. That is, we need to compute the following for each guess of γ.

$$\mu(\gamma) = \inf_{K(s) \in (RH^\infty)^{mxr}} \| [N(s) + \phi(s)K(s)] M(s,\gamma)^{-1} \|_\infty \qquad (4\text{-}3a)$$

where

$$\gamma^2 I - S(-s)^T S(s) = M(-s,\gamma)^T M(s,\gamma) \qquad (4\text{-}3b)$$

The computation of $\mu(\gamma)$ by using the method of Chang and Pearson [8] is described as follows. Suppose all the zeros $\{s_k, k=1,2,....,n\}$ of $\phi(s)$ are distinct and define W^k as follows

$$W^k = N(s_k)\, M(s_k,\gamma)^{-1} \qquad k=1,2,....,n \qquad (4\text{-}4)$$

Without loss of generality, we assume that W^k is square, i.e., m=r. If W^k is not square, some zero row (or column) vectors can be added so that W^k is square [8]. Define

$$W = \text{blockdiag}\,(W^1, W^2, \ldots, W^n) \qquad (4\text{-}5)$$

where $W^k \in \mathbb{C}^{rxr}$, k=1,2,...,n, are given by (4-4). $\mathbb{C}^{rxr}$ is the set of rxr matrices whose entries are complex numbers. Define

$$V = [V_{ij}] \qquad i,j = 1, 2, ..., n \tag{4-6}$$

with

$$V_{ij} = \frac{1}{s_i + s_j^*} I_r$$

Then the optimal norm $\mu(\gamma)$ can be computed by using the following theorem.

Theorem 4-1: [8]

The optimal norm $\mu(\gamma)$ is the square root of the maximum λ such that

$$\det (W V W^* - \lambda V) = 0 \tag{4-7}$$

Corollary 4-1: [8]

Define $V = R R^*$. Then the optimal norm $\mu(\gamma)$ is the maximum singular value of $R^{-1} W R$.

Now the computation of $\mu(\gamma)$ in (4-3a) can be reduced to a generalized eigenvalue problem or a singular value problem. To use Theorem 4-1 or Corollary 4-1, we need the following information, the zeros $\{ s_k, k=1,2,...,n\}$ of $\phi(s)$, and the constant matrices of $N(s_k)$ and $M(s_k,\gamma)$, k=1,2,....,n. Recall that $N(s) = \phi(s)I_m R_1(s) = R_1(s)\phi(s)I_r$. For N(s) to be stable, the unstable poles of $R_1(s)$ should be completely cancelled out by the zeros of $\phi(s)I_m$ or $\phi(s)I_r$. The pole-zero cancellation can be done numerically by minimal realization techniques. Nevertheless, the unstable poles of $R_1(s)$ can be completely removed from N(s) by a much simpler method without using any numerical minimal realization technique. In this section an elegant state-space approach for the evaluation of $N(s_k)$ will be presented.

If we use [13,14] to compute the spectral factorization of $\gamma^2 I - R_2^* R_2$, we need to find a stable state-space realization of $S(s) = U(s)R_2(s)$ where U(s) is an inner matrix. Doyle [13] and Chu [14] had a method for constructing stable S(s) based on a minimal realization of $R_2(s)$. An alternate method is proposed in the section. The alternate method does not require a minimal realization of $R_2(s)$ and only a Lyapunov equation is involved.

First of all, we will summarize the procedures of evaluating $N(s_k)$, k=1,2....,n, where

$$N(s) = \phi(s)I_m R_1(s) = R_1(s)\phi(s)I_r \tag{4-8}$$

Recall that $R_1(s) \in (RL^\infty)^{mxr}$ is completely unstable and a state-space realization of $R_1(s)$ is given as

$$R_1(s) = \left[\begin{array}{c|c} -A^T & B \\ \hline C & D \end{array}\right] \tag{4-9}$$

where $A \in \mathbb{R}^{nxn}$ is stable, $B \in \mathbb{R}^{nxr}$, $C \in \mathbb{R}^{mxn}$, and $D \in \mathbb{R}^{mxr}$. $\phi(s)$ is an inner function whose zeros $\{ s_k, k=1,2,...,n \}$ are the poles of $R_1(s)$, or the eigenvalues of $-A^T$. Note that these zeros are all in the open right half plane.

It can be easily shown that the D matrix in the realization of $R_1(s)$ has no effect on the evaluation of $N(s_k)$, k=1,2,...,n, and therefore we have

$$N(s_k) = R_1(s)\,\phi(s)\,|_{S=s_k} = R_{10}(s)\,\phi(s)\,|_{S=s_k} \tag{4-10}$$

where the realization of $R_{10}(s)$ is the same as that of $R_1(s)$ except D = 0, i.e.,

$$R_{10}(s) = \left[\begin{array}{c|c} -A^T & B \\ \hline C & 0 \end{array}\right] \tag{4-11}$$

Before we plug $s = s_k$, k=1,2,...,n, into $R_{10}(s)\phi(s)$, all the poles of $R_{10}(s)$ should have been cancelled out by the zeros of $\phi(s)I_m$ (or $\phi(s)I_r$). The pole-zero cancellation and then the evaluation of $N(s_k)$ can be easily done by using either Theorem 4-2 or Theorem 4-3.

Theorem 4-2: [2]

Suppose $b_1, b_2, \ldots, b_r$ are columns of B, $c_1, c_2, \ldots, c_m$ are rows of C. Then

$$N(s_k) = [\, n_{ij}(s_k) \,] \qquad \begin{array}{l} k = 1,2,...,n \\ i = 1,2,....,m \\ j = 1,2,...,r \end{array} \tag{4-12}$$

with

$$n_{ij}(s_k) = (c_i X_j (s_k I_n - A)^{-1} X_j^{\dagger} b_j) \varphi_j(s_k) \tag{4-13}$$

where $X_j = X_j^T$ satisfies the following Lyapunov equation

$$A^T X_j + X_j A + b_j b_j^T = 0 \qquad j = 1,2,...,r \tag{4-14}$$

and $X_j^{\dagger}$ is the pseudo inverse of X_j and $\varphi_j(s)$ is an inner function whose poles are the stable eigenvalues of $A + X_j^{\dagger} b_j b_j^T$. Note that $\varphi_j(s) = 1$ if X_j is invertible.

Remark 4.1: If X_j is invertible, it is easy to see that the eigenvalues of $A + X_j^{-1} b_j b_j^T$ are those of $-A^T$, i.e., $A + X_j^{-1} b_j b_j^T$ has no stable eigenvalues and therefore $\varphi_j(s) = 1$.

Remark 4.2: If X_j is singular, $A + X_j^{\dagger} b_j b_j^T$ has some stable eigenvalues, say, $-z_1$, $-z_2$, . . , and $-z_p$, then the inner function $\varphi_j(s)$ can be constructed as

$$\varphi_j(s) = \frac{s - z_1}{s + z_1} \frac{s - z_2}{s + z_2} \cdots \frac{s - z_p}{s + z_p} \tag{4-15}$$

and $\varphi_j(s_k)$ can be easily evaluated.

To compute $N(s_k)$, k=1,2,...,n, by using Theorem 4-2, we need to solve r Lyapunov equations. In the case of $m \geq r$, Theorem 4-2 is the right choice. On the other hand, i.e., $m < r$, we recommend Theorem 4-3 by which only m Lyapunov equations are involved.

Theorem 4-3: [2]

$R_1(s)$, $\phi(s)$, b_j, and c_i, $i = 1,2,...,m$, $j = 1,2,...,r$, are defined in Theorem 4-2. Then

$$N(s_k) = [\, n_{ij}(s_k) \,] \qquad \begin{matrix} k = 1,2,...,n \\ i = 1,2,...,m \\ j = 1,2,...,r \end{matrix} \tag{4-16}$$

with

$$n_{ij}(s_k) = \varphi_i(s_k) \{ c_i Y_i^{\dagger} (s_k I_n - A)^{-1} Y_i b_j \} \tag{4-17}$$

where $Y_i = Y_i^T$ satisfies the following Lyapunov equation

$$A Y_i + Y_i A^T + c_i^T c_i = 0 \qquad i = 1,2,...,m \tag{4-18}$$

and $Y_i^{\dagger}$ is the pseudo inverse of Y_i and $\varphi_i(s)$ is an inner function whose poles are the stable eigenvalues of $A + c_i^T c_i Y_i^{\dagger}$. Note that $\varphi_i(s) = 1$ if Y_i is invertible.

The second problem to be considered in the section is the construction of a stable

$$S(s) = U(s) R_2(s) \tag{4-19}$$

where U(s) is an inner matrix to be determined and $R_2(s) \in (RL^{\infty})^{q \times r}$ is completely unstable and a state-space realization of $R_2(s)$ is given as

$$R_2(s) = \left[\begin{array}{c|c} -A_2^T & B_2 \\ \hline C_2 & D_2 \end{array}\right] \tag{4-20}$$

where $A_2 \in \mathbb{R}^{n\times n}$ is stable, $B_2 \in \mathbb{R}^{n\times r}$, $C_2 \in \mathbb{R}^{q\times n}$, and $D_2 \in \mathbb{R}^{q\times r}$. This problem was solved by Doyle [13] and Chu [14] based on a minimal realization of $R_2(s)$ and a solution of a Riccati equation. In the following, a new method is proposed. The new method does not require a minimal realization of $R_2(s)$ and only a Lyapunov equation is involved. The new method of constructing a stable S(s) is summarized in the following theorem.

Theorem 4-4: [2]

An inner matrix U(s) can be found such that S(s) is stable and S(s) can be realized as follows

$$S(s) = U(s)\,R_2(s) = \left[\begin{array}{c|c} A_2 & -YB_2 + C_2^T D_2 \\ \hline -C_2 Y^\dagger & D_2 \end{array}\right] \tag{4-21}$$

where $Y = Y^T \in \mathbb{R}^{n\times n}$ is a solution of the following Lyapunov equation

$$A_2 Y + Y A_2^T + C_2^T C_2 = 0 \tag{4-22}$$

and $Y^\dagger$ is the pseudo inverse of Y.

V. FAST ITERATIVE COMPUTATION OF OPTIMAL TWO-BLOCK H^∞ NORM

In (2-15), if $[\,R_{12}^T \; R_{22}^T\,] = 0$, i.e.,

$$\hat{\Phi}(s) = \begin{bmatrix} R_{11}(s) + \hat{Z}(s) \\ R_{21}(s) \end{bmatrix} \tag{5-1}$$

then we have the <u>two-block H^∞ optimization problem</u> described as follows,

> Find a $\hat{Z}(s) \in (RH^\infty)^{m\times r}$ such that $\|\hat{\Phi}\|_\infty$ is minimized
> where $\hat{\Phi}$ is given by (5-1). (5-2)

The two-block H^∞ optimization problem arises in the optimal disturbance attenuation with control weighting, the minimization of a weighted sum involving both the sensitivity function and its complement, and any control problems with more controlled outputs than control inputs. The two-

block H^∞ optimization problem usually is solved in two steps. The first step is the computation of the optimal H^∞ norm $\| \hat{\Phi} \|_\infty$ which is the most time-consuming job. The second step is the construction of an optimal $\hat{Z}(s)$ and then an optimal controller.

In [3], a very fast iterative algorithm was developed for the computation of the optimal two-block H^∞ norm, i.e., the computation of

$$\inf_{K(s) \in (RH^\infty)^{mxr}} \left\| \begin{bmatrix} N(s) + \phi(s)\, K(s) \\ S(s) \end{bmatrix} \right\|_\infty := \gamma_o \tag{5-3}$$

N(s) and S(s) are in $(RH^\infty)^{mxr}$ and $(RH^\infty)^{qxr}$ respectively, and $\phi(s)$ is an inner function, i.e., $\phi(s) \in RH^\infty$ and $|\phi(j\omega)| = 1$ for all ω. The relation between N(s), S(s), $\phi(s)$, and $R_{11}(s)$, $R_{21}(s)$ will be given in the next section.

In the two-block H^∞ optimization problem, the most computationally demanding work is the computation of the optimal H^∞ norm γ_o. Define

$$\mu(\gamma) = \inf_{K(s) \in (RH^\infty)^{mxr}} \| [N(s) + \phi(s)K(s)]\, M(s,\gamma)^{-1} \|_\infty \tag{5-4a}$$

where

$$\gamma^2 I - S(-s)^T S(s) = M(-s,\gamma)^T M(s,\gamma) \tag{5-4b}$$

Then the problem of computing γ_o can be considered as that of finding a γ such that $\mu(\gamma)$ equals to 1.

In γ-iterative approach, an initial guess of the optimal norm γ_o, say γ, is made and the computations of the spectral factorization (5-4b) and the one-block optimal H^∞ norm computation (5-4a) are performed to determine what the next guess shall be. The next guess of γ is determined by the value of $\mu(\gamma)$. This guess and computation loop is repeated until $\mu(\gamma) = 1$. Numerically, this computation ends when $|\mu(\gamma) - 1| < \varepsilon$, where ε is a very small number. This search scheme works all right if $|\mu(\gamma) - 1| < \varepsilon$ implies that $|\gamma - \gamma_o| < \varepsilon$, i.e., the absolute value of the slope of $\mu(\gamma)$ at γ is large enough. On the other hand, if the absolute value of the slope is small, we may have $|\gamma - \gamma_o| >> \varepsilon$. In other words, the accuracy of γ may be very poor although the accuracy of $\mu(\gamma)$ is good.

A simple bisection search scheme can be used. Starting from a lower bound γ_L and an upper bound γ_U, one can define $\gamma = (\gamma_L + \gamma_U)/2$ and evaluate $\mu(\gamma)$. If $\mu(\gamma) > 1$, then γ_L is updated by γ. Otherwise, γ_U is replaced by γ. This process is repeated until the gap between γ_L and γ_U is small

enough. This search scheme is straight forward, but the convergence rate is slow. Say, the initial gap between γ_L and γ_U is 1. To reduce the gap to be less than 10^{-15}, the number of iterations required is

$$n > 15 / (\log_{10} 2) = 49.8$$

In [3] a very fast iterative algorithm was developed for the computation of the optimal H^∞ norm γ_o in (5-3). To achieve this objective, three tasks have been done. The first task is in finding better initial lower and upper bounds of γ_o. Here "initial lower and upper bounds" means the lower and upper bounds which can be easily computed without doing any spectral factorization. The second task is to reducing the computation amount in each iteration. The last, but the most important task is to find a fast search scheme by which the number of iterations is greatly reduced.

In [3], we presented an easy way of computing a new initial lower bound and showed some useful properties of $\mu(\gamma)$ which had just been discovered. Together with the well-known properties of $\mu(\gamma)$, these discoveries are of great help in developing a sophisticated and fast search scheme for γ_o. The initial lower and upper bounds are denoted by γ_1 and γ_2 respectively.

Before entering the iteration loop, we compute $\mu(\gamma_1)$, and $\mu(\gamma_2)$. With the knowledge of γ_1, $\mu(\gamma_1)$, γ_2, and $\mu(\gamma_2)$, we can generate three new upper-bound candidates and two new lower-bound candidates very easily. A new greatest lower bound, denoted by γ_3, and a new least upper bound, denoted by γ_4, can be found.

Now we are starting the iteration loop with the input data: γ_1, $\mu(\gamma_1)$, γ_2, $\mu(\gamma_2)$, γ_3, and γ_4. Let $\gamma = (\gamma_3 + \gamma_4)/2$, and evaluate $\mu(\gamma)$. If $\mu(\gamma)$ is exactly equal to 1, then γ is the optimal norm γ_o. Otherwise, $\mu(\gamma)$ may be greater or less than 1. In either case, we can easily generate two new lower-bound candidates and two new upper-bound candidates by using the information of γ, $\mu(\gamma)$, γ_1, $\mu(\gamma_1)$, γ_2, and $\mu(\gamma_2)$. The bounds γ_3, and γ_4 can be updated by the new greatest lower bound and the new least upper bound respectively. Depending on the sign of $\mu(\gamma)$-1, either the pair $\{\gamma_1, \mu(\gamma_1)\}$ or the pair $\{\gamma_2, \mu(\gamma_2)\}$ will be updated by $\{\gamma, \mu(\gamma)\}$. The iteration loop is repeated until the difference $\gamma_4 - \gamma_3$ is small enough.

The search scheme described above is much faster than any other existing γ-iteration search scheme. We only require one evaluation of $\mu(\gamma)$ per iteration. In each iteration, two new lower-bound

and two new upper-bound candidates are generated. In several examples we have encountered, the gap of lower and upper bounds can be reduced from 1 to 10^{-15} in only 2 or 3 iterations.

VI. OPTIMAL H^∞ NORM COMPUTATION FOR MIMO SYSTEMS WITH MULTIPLE RHP ZEROS

In the design of optimal H^∞ controllers, we need to compute the following optimal H^∞ norm

$$\mu_o := \inf \{ \| H(s) + \phi(s)K(s) \|_\infty , \ K(s) \text{ is stable } \} \tag{6-1}$$

where $H(s) \in (RH^\infty)^{mxr}$ and $\phi(s)$ is an inner function, i.e., $\phi(s) \in RH^\infty$ and $|\phi(j\omega)| = 1$ for all ω. $(RH^\infty)^{mxr}$ is the set of mxr rational matrices with real coefficients and with entries in H^∞, which is the space of all functions that are analytic and bounded in the right half plane. The norm $\|\Phi\|_\infty$ is defined as follows,

$$\|\Phi\|_\infty := \sup_\omega \bar{\sigma}[\Phi(j\omega)] \tag{6-2}$$

where $\bar{\sigma}[\Phi(j\omega)]$ denotes the largest singular value of $\Phi(j\omega)$.

There are two major approaches for the computation of the optimal H^∞ norm in (1). One is the H^∞ interpolation method [6-8] and the other is the Hankel approximation method [9,10]. Francis and Zames [6] solved the SISO case of the problem in an elegant way by using Sarason's H^∞ interpolation theory [34]. Chang and Pearson [8] applied the Sarason's theory successfully to the MIMO case and simplified the problem to a simple eigenvalue or singular value computation. Francis, Helton, and Zames [7] solved the MIMO case of the problem by using Ball and Helton's theory [35].

Glover [10] provided an easy way to construct an optimal K(s). If it is required to construct an optimal stable K(s) which minimizes $\|H(s)+\phi(s)K(s)\|_\infty$, Glover's method is recommended. If we are only interested in computing the optimal H^∞ norm μ_o, Chang and Pearson's method is an easy way to go. The two or four block H^∞ optimization problems usually are solved by iterative approaches [19-21,3] and in each iteration we need to compute the minimal H^∞ norm without constructing K(s).

The only drawback of the Chang and Pearson's method is the restriction that the zeros of $\phi(s)$ should be distinct. In this section, the restriction will be removed and we will see that the MIMO case

of problem (6-1) with multiple RHP zeros can be also reduced to a simple eigenvalue or singular value computation problem.

Without loss of generality [8], we can assume that H(s) is a square matrix, i.e., m=r. Suppose the zeros of $\phi(s)$ are s_i with multiplicity m_i, i=1,2,...,n. In the following, we will define two constant matrices V and W. The optimal H^∞ norm μ_o can be easily obtained through a simple eigenvalue or singular value computation which involves the matrices V and W. V is defined as

$$V = [\, V_{pq} \,] \tag{6-3a}$$

with

$$V_{pq}(i,j) = \frac{(i+j-2)!}{(i-1)!\,(j-1)!} \frac{I_r}{(s_p + s_q^*)^{i+j-1}} \tag{6-3b}$$

$$p,q = 1,2,...n. \qquad i=1,2,...,m_p. \qquad j=1,2,.....,m_q.$$

where I_r is an rxr identity matrix, $V_{pq}(i,j)$ is the i-j block of the matrix V_{pq} and V_{pq} is the p-q block of the matrix V. The matrix W is defined by

$$W = \text{blockdiag}\,(\, W_{11}, W_{22}, \ldots, W_{nn} \,) \tag{6-4a}$$

where W_{kk} is a lower block triangular Toeplitz matrix which is defined by the following first column elements

$$W_{kk}(i,1) = \frac{(-1)^{i-1}}{(i-1)!} H^{(i-1)}(s_k) \tag{6-4b}$$

$$i = 1,2,...,m_k. \qquad k = 1,2,...,n.$$

where $H^{(i-1)}(s)$ is the (i-1)-th derivative of H(s). Furthermore, if a state space realization of H(s) is given as (A_H, B_H, C_H, D_H), i.e.,

$$H(s) = C_H (sI - A_H)^{-1} B_H + D_H \tag{6-5}$$

we have

$$\begin{aligned} W_{kk}(1,1) &= C_H (s_k I - A_H)^{-1} B_H + D_H \\ W_{kk}(i,1) &= C_H (s_k I - A_H)^{-i} B_H \end{aligned} \tag{6-6}$$

$$i = 2,3,...,m_k. \qquad k = 1,2,...,n.$$

Then we have the following main theorem.

Theorem 6-1:

The optimal H^∞ norm μ_o is the square root of the maximum λ such that

$$\det (WVW^* - \lambda V) = 0 \qquad (6\text{-}7)$$

where V and W are given in (6-3) and (6-4), respectively.

Corollary 6-1:

V is Hermitian and positive definite and can be factorized as $V = UU^*$. Then μ_o is the maximum singular value of $U^{-1}WU$.

Now, the problem of computing the optimal H^∞ norm μ_o is reduced to a generalized eigenvalue problem or a singular value problem. There are many existing computer packages available, e.g., EISPACK, LINPACK, MATLAB, MATRIXx, etc. for these reduced problems.

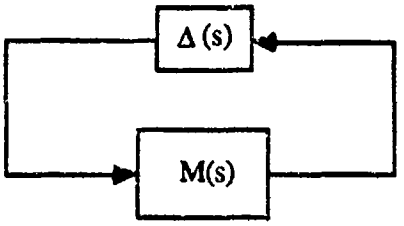
VII. CONCLUSION AND FURTHER RESEARCH

Several significant contributions have been made in the research period supported by AFOSR/UES Mini Grant. These contributions are summarized in the previous sections of the report and in following papers: [1], [2], [3], and [4].

In the next few years, we will emphasize our research in the following three aspects: (1) Continue the theoretical research in robust multivariable control systems and develop a better design methodology. (2) Apply the design methodology to flight control systems, flexible structured robot control systems, and other control systems related to the aerospace program, and (3) Develop a user friendly computer-aided-design package based on the design methodology.

Two robustness issues will be considered. One is the command/disturbance input uncertainties and the other is the plant uncertainties. The controller is to be designed such that the closed-loop system remains stable for all possible plant perturbations and the error response is limited to a desired region for every disturbance and command in the prescribed set under all possible plant perturbations.

Most of plant uncertainties, structured or unstructured, unmodeled dynamics or parametric perturbations, can be represented by the following block diagram,



Standard structure for a perturbed closed-loop system.

where $\Delta(s)$ = block diag { $\Delta_1(s)$, $\Delta_2(s)$, ..., $\Delta_m(s)$ } and M(s) is the nominal closed-loop system which includes the nominal plant and the stabilizing controller. Doyle define the structured singular value (SSV) as the inverse of the minimal size of Δ_i for all i which destabilizes the system. The SSV is the most unconservative scalar stability-margin measure we ever have for multivariable systems.

Algorithms [32-36] to compute the SSV are available only for the cases with the number of perturbation blocks less than or equal to three. This computation problem for the cases which have more than three perturbation blocks is still a hot research problem.

One important special case of plant uncertainties is the real parametric perturbation. In this case the perturbation matrix $\Delta(s)$ is a real block diagonal matrix. The SSV defined for this case is called real SSV. Algorithms of computing the real SSV [37,38] are also only available for the cases with the number of perturbation blocks less than or equal to three.

For a stable nominal closed-loop system it is natural to ask how stable the system is, i.e., what limit the system can perturb without causing instability. Stability-margin measures like gain margin, phase margin, and the SSV, etc. are developed for this purpose. These measures are either simple nonnegative real scalars or nonnegative real scalar functions (function of frequency). By any of these stability-margin measures it is impossible to describe the boundary of the largest perturbation space in which the system is stable. For example, let's consider a nominal closed-loop system M(s) with two perturbed blocks Δ_1 and Δ_2. Say, the perturbed system remains stable whenever $\|\Delta_1\|$ and $\|\Delta_2\|$ (the sizes of Δ_1 and Δ_2) are less than 1 and 10 respectively. In this example the SSV of the system is unity. Unity SSV means that the perturbed system remains stable if all perturbed blocks have sizes less than one. The SSV of the system tells us nothing about how far Δ_2 can perturb without causing instability.

It is more practical to look at the robust stability problem from the other direction. Usually it is possible for the designer to define the set of all possible perturbed plants according to the modeling approximations, component tolerances, and the changing environment conditions, etc. The robust stability problem can be considered as a problem of finding a realizable controller such that the closed-loop system is stable for any perturbed plant in the set. For the same example of the previous paragraph, we only need to find a controller such that the closed-loop system is stable for all perturbed plants with that two-block structure satisfying $\|\Delta_1\| \leq 1$, and $\|\Delta_2\| \leq 10$. If we use SSV as the stability-margin constraint, then the problem would be to make the closed-loop system stable for all perturbed plants satisfying $\|\Delta_1\| \leq 1$, and $\|\Delta_2\| \leq 1$. Obviously, the latter approach is unnecessarily conservative.

Once an optimal controller is synthesized, we may need to reduce the order of the controller for easy implementation. There are plenty of model reduction techniques available. However, it is still an important issue to find a model reduction scheme which reduces the order of the controller but its influence on the closed-loop performance and stability is minimized.

LIST OF REFERENCES

[1] B-C Chang, "Size Reduction in Four-Block H^∞ Formulation", Proceedings of the 26th Conference on Decision and Control, Dec. 1987.

[2] B-C Chang, "A Stable State-Space Realization in the Formulation of H^∞ Norm Computation", IEEE Trans. on A-C, AC-32, pp.811-815, Sept. 1987.

[3] B-C Chang, S. S. Banda, T. E. McQuade, "Fast Iterative Algorithm for H^∞ Optimization Problems", Proceedings of the American Control Conference, Minneapolis, MN, June 1987.

[4] B-C Chang and S. S. Banda, "Optimal H^∞ Norm Computation for MIMO Systems with Multiple RHP Zeros", Submitted to the 1988 American Control Conference.

[5] G. Zames, "Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Seminorms, and Approximate Inverses", IEEE Trans. A-C, AC-26, pp. 301-320, 1981.

[6] B. A. Francis and G. Zames, "On H^∞-Optimal Sensitivity Theory for SISO Feedback Systems", IEEE Trans. A-C, AC-29, pp.9-16, Jan. 1984.

[7] B. A. Francis, J. W. Helton, and G. Zames, "H^∞-Optimal Feedback Controllers for Linear Multivariable Systems", IEEE Trans. A-C, AC-29, pp.888-900, Oct. 1984.

[8] B-C Chang and J. B. Pearson, "Optimal Disturbance Reduction in Linear Multivariable Systems", IEEE Trans. A-C, AC-29, pp.880-887, Oct. 1984.

[9] M. G. Safonov and M. S. Verma, "L^∞-sensitivity Optimization via Hankel-norm Approximation", Proceedings of American Control Conference, San Francisco, June 1983.

[10] K. Glover, "All Optimal Hankel-norm Approximations of Linear Multivariable Systems and Their L^∞-error Bounds", International Journal of Control, Vol. 39, No. 6, pp.1115-1193, 1984.

[11] A. C. Allison and N. J. Young, "Numerical Algorithms for the Nevanlinna-Pick Problem", Numerische Mathematik, Vol. 42, pp.125-145, 1983.

[12] J. Doyle, "Synthesis of Robust Controllers and Filters", Proceedings of 22nd Conference on Decision and Control, Dec. 1983.

[13] J. Doyle, Lectures Notes, ONR/Honeywell Workshop on Advances in Multivariable Control, Minneapolis, Minnesota, Oct. 1984.

[14] C-C Chu, H_∞ Optimization and Robust Multivariable Control, Ph.D. Dissertation, University of Minnesota, Sept. 1985.

[15] H. Kimura, "Robust Stabilization for a Class of Transfer Functions", IEEE Trans. A-C, AC-29, pp.788-793, Sept. 1984.

[16] H. Kwakernaak, "Minimax Frequency Domain Performance and Robustness Optimization of Linear Feedback Systems", IEEE Trans. A-C, AC-30, pp.994-1004, Oct. 1985.

[17] B. A. Francis, Notes On H^∞-Optimal Linear Feedback Systems, Lectures presented in the Division of Automatic Control, Department of Electrical Engineering, Linköping University, Aug. 1983.

[18] A. [illegible]ch and B. A. Francis, "Uniformly Optimal Control of Linear Feedback Systems", Automatica, vol.21, No.5, pp.563-574, 1985.

[19] B-C Chang and J. B. Pearson, "Iterative Computation of Minimal H^∞ Norm", Proceedings of 24th Conference on Decision and Control, Dec. 1985.

[20] C-C Chu, J. Doyle, and E. B. Lee, "The General Distance Problem in H_∞ Optimal Control Theory", International Journal of Control, Vol. 44, No. 2, pp.565-596, 1986.

[21] E. A. Jonckheere and J. C. Juang, "Toeplitz + Hankel" Structures in H_∞ Design and The Avoidance of the ϵ-Iteration", Proceeding of American Control Conference, Seattle, June 1986.

[22] D. J. N. Limebeer and Y. S. Hung, "An Analysis of the Pole-zero Cancellations in H^∞ Optimal Control Problems of the First Kind", Preprint, Dept. of E. E., Imperial College, London.

[23] I. Postlethwaite, D. W. Gu, and S. D. O'Young, "Some Computational Results on Size Reduction in H_∞ Design", Preprint, Dept. of Engineering Science, University of Oxford, U.K.

[24] H. Kwakernaak and R. Sivan, Linear Optimal Control Systems, John Wiley & Sons, Inc., 1972.

[25] D. C. Youla, J. J. Bongiorno, and H. A. Jabr, "Modern Wiener-Hopf Design of Optimal Controllers, Part II: The Multivariable Case", IEEE Trans. A-C, AC-21, pp.319-338, 1976.

[26] C.A. Desoer, R.W. Liu, J. Murray, and R. Saeks, "Feedback System Design: The Fractional Representation Approach", IEEE Trans. A-C, AC-25, pp.399-412, 1980.

[27] C. N. Nett, C. A. Jacobson and M. J. Balas, "A Connection Between State-space and Doubly Coprime Fractional Representations", IEEE Trans. A-C, AC-29, pp. 831-832, 1984.

[28] G. Bengtsson, "Output Regulation and Internal Models - A Frequency Domain Approach", Automatica, 13, pp. 333-345, 1977.

[29] L. Pernebo, "An Algebraic Theory for the Design of Controllers for Linear Multivariable Systems - Part II: Feedback Realizations and Feedback Design", IEEE Trans. A-C, AC-26, pp.183-193, 1981.

[30] L. Cheng and J. B. Pearson, "Synthesis of Linear Multivariable Regulators", IEEE Trans. A-C. AC-26, pp. 194-202, 1981.

[31] C. A. Desoer and M. Vidyasagar, Feedback Systems: Input-Output Properties, Academic Press, 1975.

[32] J. Doyle, "Analysis of Feedback Systems with Structured Uncertainties", IEE PROC. Vol.129 Pt.D. No.6, pp. 242-250, 1982.

[33] J. Doyle, "Structured Uncertainty in Control System Design", Proceedings of the 24th Conference on Decision and Control, Dec. 1985.

[34] M. K. H. Fan and A. L. Tits, "Characterization and Efficient Computation of the Structured Singular Value", IEEE Trans. A-C. AC-31, pp. 734-743, Aug. 1986.

[35] M. K. H. Fan and A. L. Tits, "Geometric Aspects in the Computation of the Structured Singular Value", Proceedings of the 1986 American Control Conference.

[36] M. G. Safonov and J. Doyle, "Minimizing Conservativeness of Robustness Singular Values", in S. G. Tzefestas (ed.), Multivariable Control, D. Reidel Publishing Company, 1984.

[37] B. G. Morton and R. M. McAfoos, "A Mu-Test for Real Parameter Variations", Proceedings of the 1985 American Control Conference.

[38] R. D. Jones, "Structured Singular Value Analysis for Real Parameter Variations", 1987 AIAA Conference in Guidance and Control.

[39] J. C. Doyle and G. Stein, "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis", IEEE Trans. A-C, AC-26, pp.4-16, Feb. 1981.

[40] N. A. Lehtomaki, N. R. Sandell, and M. Athans, "Robustness Results in Linear-Quadratic-Gaussian Based Multivariable Control Design", IEEE Trans. A-C, AC-26, pp.75-92, Feb. 1981.

Appendices can be obtained from

Universal Energy Systems, Inc.

Final Report

VISUALIZATION OF HYDROCARBON JET DIFFUSION FLAMES

Submitted to

Mini Grant Program
Universal Energy Systems, Inc.
4401 Dayton-Xenia Road
Dayton, Ohio 45432

Contract Number: F49620-85-C-0013/SB5851-0360

Subcontract Number: S-760-6MG-113

Report Period: January 1, 1987 to December 31, 1987

Prepared By

L. D. Chen

Associate Professor
Department of Mechanical Engineering
The University of Iowa
Iowa City, Iowa 52242

January 1988

Abstract

Recently, Buckmaster and Peters showed that the flame oscillation is a result of buoyancy driven Kelvin-Helmholtz instability. In this report, the flow structure due to buoyancy driven Kelvin-Helmholtz instability and its relations with respect to the flame oscillation were identified. The oscillation frequency as well as the convective velocity of buoyancy driven toroidal vortices were characterized for ethylene, methane, and propane diffusion flames over a wide range of test conditions. The toroidal vortices are shown to exert significant stretching on the flame surface, capable of quenching the flame surface at low fuel flow rates. The toroidal vortices are also found responsible for the pairing or merging of the flame bulge.

NOMENCLATURE

Symbol	*Description*
Do	burner diameter.
Q	volumetric flow rate.
Re	burner exit Reynolds number based on Do, Uo and kinematic viscosity of the fuel gas at 298 K.
Ua	annulus air velocity.
Uo	burner exit velocity ($= 4Q/\pi Do^2$).
z	vertical distance above the burner exit plane.

ACKNOWLEDGEMENT

The technical assistance from R. L. Britton, C. Reaves, L. P. Goss, D. D. Trump, J. P. Seaba and P. S. Tschen is acknowledged. The support from AFWAL/APL/POSF in conducting the research is also acknowledged. Special thanks are due to Dr. W. M. Roquemore for enlightening discussion and collaboration.

I. INTRODUCTION

Low frequency flame oscillations, also known as the flame flicker[1] and flare,[2] have been observed for some time. The flicker frequency observed in nonpremixed and partially premixed flames is typically in the range of 10 to 20 Hz, e.g., see Chamberlin and Rose,[1] Remenyi,[2] Kimura,[3] Grant and Jones,[4] Ballantyne and Bray,[5] Yule, et. al.,[6] Becker and Liang,[7] Eickhoff,[8] Durao and Whitelaw,[9] and Buckmaster and Peters.[10] The frequency is relatively independent of the fuel type, nozzle size, and jet exit velocity. This "invariant" nature has puzzled the combustion community for years. Different hypotheses and conjectures were proposed to explain the phenomenon. It is, however, only recently shown by Buckmaster and Peters[10] that the oscillation is due to buoyancy driven Kelvin-Helmholtz instability.

Two different vortices have been identified in jet diffusion flames,[11,12,13] small roll-up vortices inside the luminous flame and large toroidal vortices outside the luminous flame. The small vortices are due to Kelvin-Helmholtz instability of the jet, similar to that observed in cold jets. The role of the inner vortices in turbulent jet diffusion flames in relation to laminar flamelets and the lift-off phenomenon were discussed in.[14] The toroidal vortices outside the luminous flame are due to buoyancy driven Kelvin-Helmholtz instability. The occurring frequency and convective velocity of the toroidal vortices, and its relation with the flame oscillation will be discussed in this paper, also discused is its impact on the flame dynamics.

The present study employs a two-dimensional visualization technique known as the Reactive Mie Scattering (RMS) in which Mie scattering TiO_2 particles were formed resulting from chemical reaction between the combustion product H_2O and the seeded gas $TiCl_4$. The RSM technique can provide detailed information on flow structures[13,15,16] and can be applied rather easily. More detailed discussion on the technique can be found in.[14,15,16]

II. EXPERIMENTAL METHODS

Mie scattering particles are formed in the flowfield resulting from mixing between fuel and combustion products, or air and combustion products. In the experiments, a small amount of $TiCl_4$ vapor was added to the fuel and air. As a result, TiO_2 particles are formed due to the chemical reaction of $TiCl_4$ and the combustion product H_2O. The theoretical stoichiometry of the reaction is as follows:

$$TiCl_4 + 2\,H_2O = TiO_2 + 4\,HCl \qquad (1)$$

The amount of $TiCl_4$ added to the flow was not metered but limited to the vapor pressure of 1.38 kPa at 298 K.[17] The reaction of $TiCl_4$ and H_2O is weakly exothermic. The effects on the flame properties, however, are not significant. The equilibrium calculations employing a NASA equilibrium code[18] yield stoichiometric temperatures of 2168 K (at mixture fraction of 0.0584), 2312 K (at mixture fraction of 0.0636), and 2210 K (at mixture fraction of 0.0584), respectively for methane, ethylene, and propane. The above calculation was made setting the vapor pressure of $TiCl_4$ at 1.38 kPa and the reactant temperature at 298 K. The corresponding stoichiometric conditions without $TiCl_4$ for methane, ethylene, and propane are 2226 K (at mixture fraction of 0.0548), 2370 K (at mixture fraction of 0.0634), and 2266 K (at mixture fraction of 0.0600), respectively. The 55-60 K decrease in stoichiometric temperatures is mainly due to the equilibrium shift as HCl and TiO_2 formed at the expense of otherwise formed combustion products, H_2O and CO_2.

The mean diameter of TiO_2 particles formed due to the reaction of $TiCl_4$ and moist room air was reported 1.1 μm.[19] The frequency response of 1 μm TiO_2 particles in gas flows was known in the range of 1 to 10 kHz.[20] The frequency response is adequate for the Mie scattering experiments at the conditions studied in this paper.

The test apparatus was described in.[14,16,21] In summary, a fuel nozzle was located at the center of a small vertical combustion tunnel pointing upward. The low velocity and low turbulence coflowing air was supplied through an annulus jet having a divergent section and a flow straightener (honeycomb and fine mesh screen). The annulus air has a diameter of 254 mm. The annulus air velocity was

maintained at 0.15 m/s for the data presented in this paper, unless specified otherwise. Dried air having a dew point of -38 °C was used. The experiments were conducted using three fuel nozzles, i. e., a contoured nozzle (Do = 10 mm) and two tube burners (Do = 11 mm and 22.5 mm, tube length of 450 mm). Commercial grade methane, ethylene, and propane were used as the fuel. The fuel and air flowrates were regulated with calibrated rotameters.

The optical arrangement for the RMS visualization was similar to.[14,16] A frequency doubled Nd:YAG laser (Quanta Ray Model DCR-2) was used as the light source (at 532 nm) for still photographs. The laser was operated at a single pulse mode (15 ns pulse output). The laser was electronically triggered when the camera shutter was opened. High speed movies were made with a high repetition rate (6000 pulse/s) copper vapor laser as the light source. The framing rate was set at 1000 fps. The cameras for still photographs and high-speed movies were positioned normal to the vertical light plane which passed through the center of the burner.

Mie scattering from TiO_2 particles illuminated by an argon ion laser (514.5 nm) or a copper vapor laser (at 510 and 578 nm) was recorded using a 1/4-m spectrometer and photomultiplier assembly similar to.[21] The spectrometer was tuned to the maximum intensity of the light source. A second photomultiplier was used to record the total radiation including emission and Mie scattering from the flame. The two photomultipliers were positioned to desired locations using an alignment He-Ne laser. The output from the photomultiplier was amplified (bandpass set at 1 kHz), recorded by an FM tape recorder (operated at a speed of 30 ips for one minute), and monitored by an oscilloscope. The signal was then analyzed by an FFT spectrum analyzer.

III. RESULTS AND DISCUSSION

Visualization of a transitional jet diffusion flame is shown in Fig. 1. The experiment was conducted with a 10 mm nozzle (Do = 10 mm) fueled with methane gas. At the burner exit, the volumetric mean velocity was maintained at 4.09 m/s. The corresponding Reynolds number (Re) was 2390, based on the fuel viscosity, nozzle diameter, and burner exit velocity. The photograph was taken with a still camera aimed at a right angle to the vertical light sheet. The vertical light sheet, as discussed earlier, was positioned across the centerline of the jet flame. The photograph thus recorded a two-dimensional image of Mie scattering superimposed with a planar projection of the three dimensional luminous flame. Also shown in the figure is a sketch of the jet flame that was obtained by tracing the luminous flame and Mie scattering images out of the photograph using a microcomputer-based tablet. The double line outside the luminous flame represents the boundary of the toroidal vortex marked by Mie scattering particles.

The visualization revealed two distinct vortical structures, i. e., large toroidal vortices outside the luminous flame (will be termed outer vortices in the discussion herein) and small roll-up vortices inside the flame (will be termed inner vortices in the discussion herein). The first identifiable outer vortex, at this particular instant, was located at 80 to 150 mm above the burner, and followed by a second vortex that was located at 150 to >200 mm above the burner. An obvious flame bugle was seen at a height (z) near 130 mm, or near the "dislocated saddle" point A. The flame bulge is probably formed due to the rotational flow inside the outer vortices. A conjectured mechanism looks like this; the toroidal vortex below the flame bulge moves the flame surface radially outward while the one above the bulge drags the flame surface inward. Thus a bulge was formed. It was also due to this conjectured view, the dislocated saddle point A was named following.[22] Inside the luminous flame, well defined roll-up vortices first appeared at a height >20 mm (or >2 Do). The inner vortices grew rapidly until a height around 40 mm was reached. Beyond that location, the size of the inner vortices remained relatively constant before a height of 80 mm, or Point B, was reached. The point B is the lower "critical" point or the boundary of the observable outer vortex. Downstream of Point B, inner vortices

seemed to be stretched before vortex pairing or coalescence was observed near the flame bulge.

The outer vortex is the flow structure due to buoyancy driven Kelvin-Helmholtz instability as concluded in.[10] The occurring frequencies of these vortices are well correlated with the flame bulge and the flame oscillation frequency. As an example, the time histories of the flame bulge, Point A, and Point B (c. f., Fig. 1) are illustrated in Fig. 2. The results were obtained by image analysis of high-speed movies (1000 fps) taken from the methane jet diffusion flame shown in Fig. 1. Periodicity in outer vortices and flame bulge are well shown by the three cycles illustrated in Fig. 2. The frequency for this particular condition is 14.3 Hz.

A two stage behavior of the flame bulge motion is seen in Fig. 2. The flame bulge first moved in parallel with Point A, until 80 ms in the first cycle, 150 ms in the second cycle, and 220 ms in the third cycle, it was then accelerated passing by that point. The acceleration of the flame bulge momentarily increased the convective velocity of Point A near the intersection of the two curves (Flame Bulge and Point A). This suggests that two inflection points with respect to time probably exist in the convective velocity of Point A; they occurred prior to and immediately after the time of the crossing. The crossing of the flame bulge through the saddle point can be seen in Fig 3. The methane diffusion flame shown in Fig. 3 was issued from a contoured nozzle with Do = 10 mm and Uo = 0.7 m/s (Re = 410). The photograph showed that a flame bulge (at z = 250 mm) accelerates passing by the saddle point located at z = 220 mm. It is due to the acceleration and the rotational motion inside the outer vortices, the flame seemed to be stretched in the axial direction. More significant stretching was observed at low fuel flow rates. At some instance, a portion of the flame can literally be quenched by the outer vortex as shown in Fig. 4. At high fuel flow rates, the flame bulge is no longer well defined; the outer vortices, however, still can be identified. For example, Fig. 5 is a visualization of a methane jet diffusion flame with a 10 mm nozzle at Uo = 11 m/s (Re = 6420). At this condition, the flame surface is not as smooth as that shown in Figs. 1 and 3. Flamelets were formed due to inner vortices colliding with otherwise smooth flame surface. The flame now experienced a high straining rate at locations where the inner vortices protrude the flame surface. The high straining rate is also experienced by the flame at locations near the burner

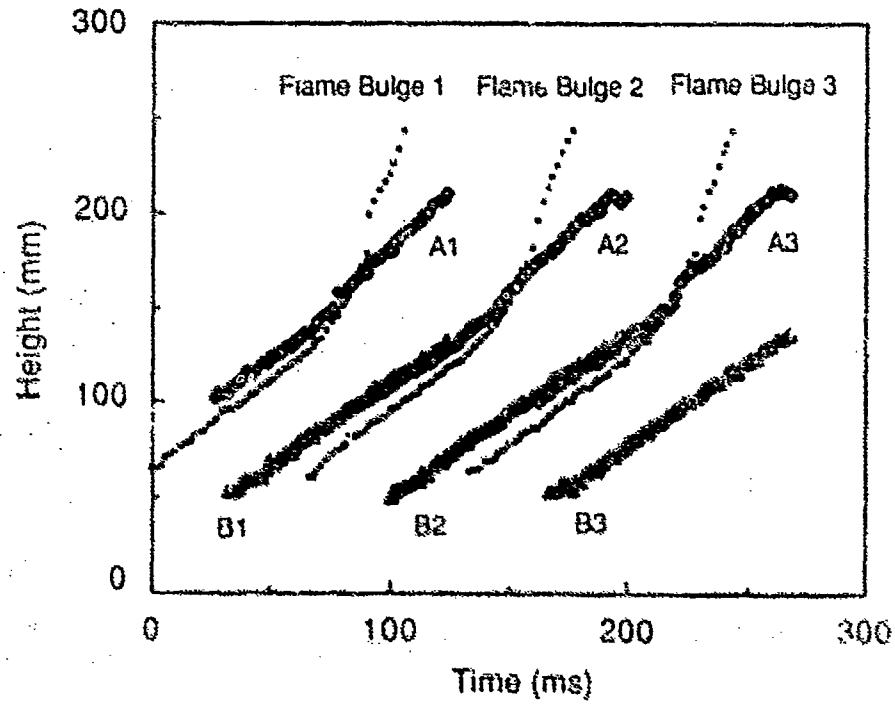


Figure 2. Time History of a Methane Jet Diffusion Flame (10 mm Nozzle, Re = 2390).

exit where high velocity gradients exist. Note that lift-off was observed for methane jet diffusion flames from the 10 mm burner at Uo = 14.9 m/s (Re = 8750).

To further illustrate the correlation between the flame bulge and the outer vortex, power spectra of the flame luminosity and Mie scattering from the outer vortex are shown in Fig. 6. The measurements were made with a tube burner (Do = 11 mm) fueled with methane gas at Re = 2370. The annulus air was maintained at 0.11 m/s for this particular condition. The data were taken at a height approximately 115 mm above the burner, using two photomultipliers. The photomultipliers were respectively aimed at the center of the flame and focused at the laser light sheet (about 30 mm away from the centerline) for flame luminosity and Mie scattering measurements. To eliminate detection of Mie scattering signals due to the flow inside the luminous flame, $TiCl_4$ seeding to the fuel jet was stopped. On the other hand, the Mie scattering detection was checked examining the signals with the laser source being blocked. Satisfactory results were obtained. A frequency of 14 Hz was recorded by the two detectors, as well as higher-order harmonics at 28 Hz and 42.5 Hz. The observation clearly showed that the flame oscillation and the outer vortex are correlated.

The results of the frequency measurements with the two techniques described above (imaging and spectrum analysis) are summarized in Fig. 7. The flame oscillation or the outer vortex occurrence frequency was found in the range of 11.4 to 17.4 Hz for the conditions summarized in Table 1-- relatively independent of the fuel type, burner geometry, and burner exit velocity, consistent with earlier studies.[1-10] A general trend, however, can be extracted from reviewing the data available to us. The frequency was found to increase as the burner velocity is increased or the coflowing air velocity is increased. The increase in coflowing air velocity will delay the inflection point in axial velocity to a downstream location. As a result, the frequency is expected to increase following the analysis by Buckmaster and Peters[10] in that a $z^{1/4}$ dependence was found. The burner geometry also exhibits a weak influence on flame oscillation frequencies. It was found[2] that long pipes generally had a higher frequency than contoured nozzles at a similar condition (Do and Re). The oscillation frequency also showed[2] a weaker dependence in Re, or Uo, a $Re^{0.07}$ dependence was found for long pipes while $Re^{0.14}$ was

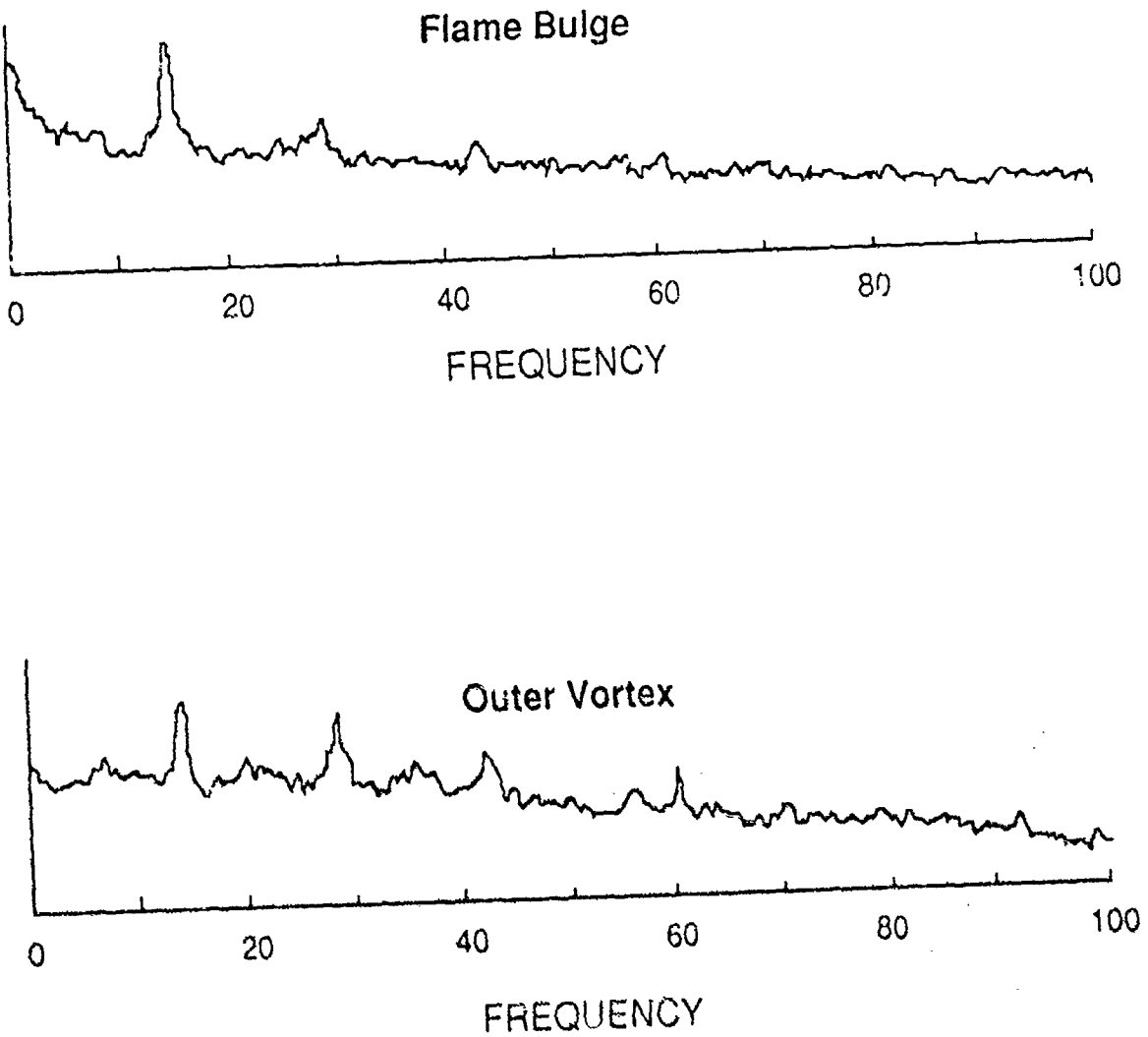


Figure 6. Spectra of a Methane Jet Diffusion Flame (11 mm Tube, Re = 2370)

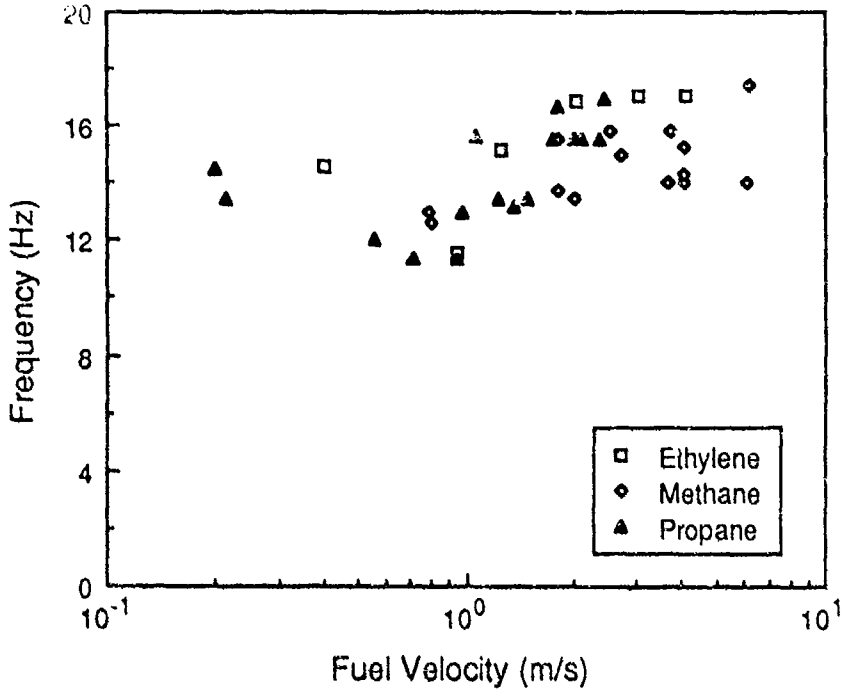


Figure 7. Frequency Data of Buoyant Jet Diffusion Flames (Test Conditions Summarized in Table 1).

Table 1. Summary of Test Conditions.

Fuel Type	Do (mm)	Burner Type	Uo (m/s)	Ua (m/s)
Ethylene	22.5	Tube	0.4	0.15
Ethylene	10	Contoured Nozzle	1.26-4.09	0.15
Methane	10	Contoured Nozzle	1.81-6.83	0.065, 0.15
Methane	11	Tube	0.79-4.02	0.1, 0.15
Methane	22.5	Tube	0.8-2.69	0.15
Propane	10	Contoured Nozzle	1.06-2.44	0.1
Propane	11	Tube	0.2-2.37	0.07, 0.1, 0.11
Propane	22.5	Tube	0.71-1.36	0.15

found for nozzles.

The vortex convective velocity extracted from image analysis is shown in Fig. 8. For most of the time, the convective velocities of Points A and B are quite constant, with A having a slightly higher velocity. The correlation coefficient of position versus time is better than 99.9 % for the data presented in Fig. 8. The slightly higher convective velocity at Point A is consistent with the observation that the outer vortex first grew axially and then radially. Growth in the radial direction is probably due to entrainment of the ambient cool air that slowed down the vortex motion allowing the flame bulge to accelerate passing by Point A. The data reported in Fig. 8 are for two burner configurations (10 mm nozzle and 11 mm tube) and three fuels (ethylene, methane and propane fuels) with the coflowing air maintained at 0.15 or 0.1 m/s. Under these conditions, the convective velocity was found in the range of 0.66 to 0.87 m/s with Point A and Point B moving at similar speeds. Point A, nevertheless, had a slightly higher velocity for most conditions. It is interesting to note that the characteristic velocity of the kinetic energy due to buoyancy forces is 0.88 m/s, based on gravitational acceleration of 9.8 m/s^2, height of 40 mm, and a density difference of 1 kg/m^3. This value is typical of the vortex convective velocity observed in the present study.

An interesting buoyancy effect on the flame dynamics is the pairing or merging of the luminous flame. The spectra of an ethylene diffusion flame (Do = 22.5 mm, Uo = 0.95 m/s, Re = 2500) are shown in Fig. 9. The photomultiplier was aimed at the center of the flame at different axial locations. Predominant frequencies of 12 and 24 Hz (fundamental and harmonic) were registered at z = 108 mm. The 12 Hz frequency is what would be normally identified as the flame oscillation frequency. At downstream locations, e.g. z = 190 and 212 mm, the strength of the harmonic frequency was increased and higher order harmonics became observable. At a location of z = 245 mm, a sub-harmonic frequency was seen. The sub-harmonic frequency of 6 Hz results from the "merging" or "pairing" of the flame bulge. The process of "merging" can be seen clearly in high-speed movies: when a flame bulge moved up and outward, its convective velocity slowed down allowing the next flame bulge to catch up. As a result, the outer flame bulge engulfed the inner one in a similar way as the vortex pairing or merging in

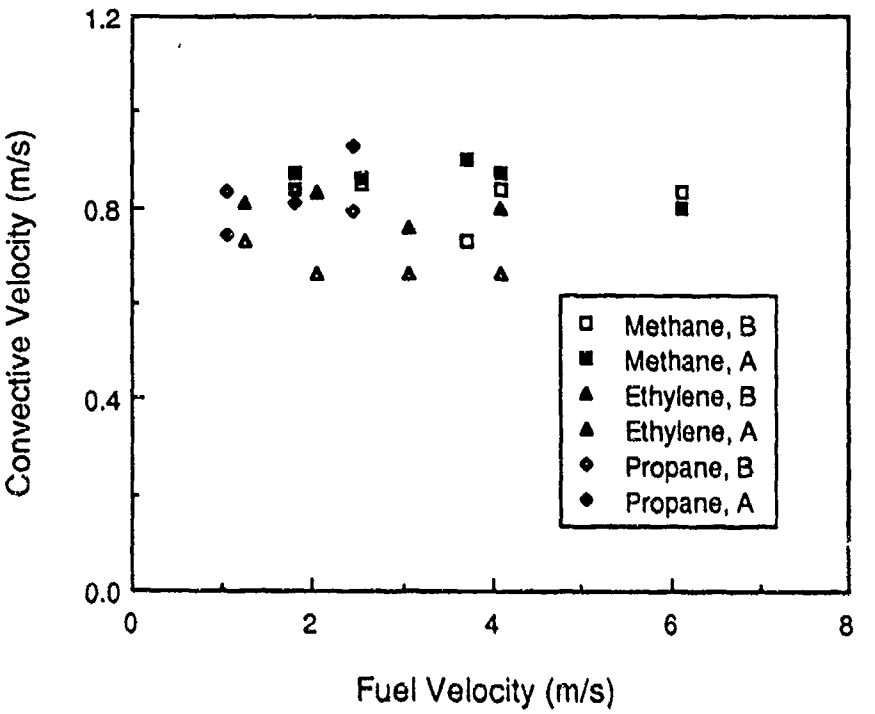


Figure 8. Convective Velocity of Buoyancy Driven Vortices.

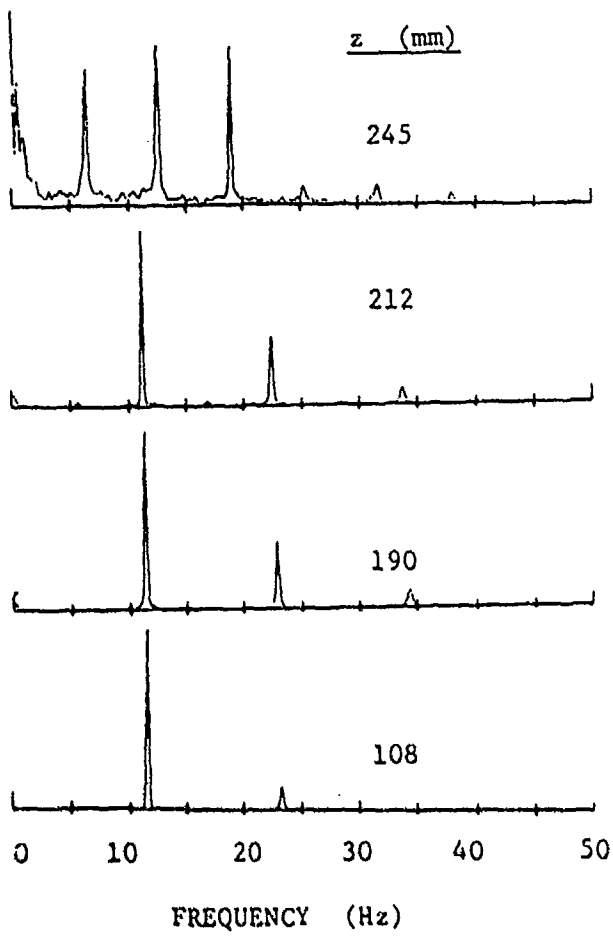


Figure 9. Spectra of Ethylene Jet Diffusion Flames (22.5 mm Tube, Re = 2500).

cold jets. The merging of the flame bulge occurred at every other flame bulge formed near the burner exit; thus a sub-harmonic frequency was registered.

Flame oscillations have been observed in jet diffusion flames at very low Reynolds number conditions as those identified in Figs. 3 and 4. A question thus arose: Are there "true" laminar diffusion flames? True laminar diffusion flames are referred to the condition when there is no oscillation in luminous flame and no fluctuation in flowfield quantities. Without fuel modifications, the necessary conditions seem to be (1) the luminous flame length being shorter than the height of the first outer vortex (approximately 30 to 40 mm), (2) the burner exit condition being laminar (i. e., low Reynolds number and clean flow condition), and (3) no disturbance from the ambient. Apparent laminar diffusion flames were observed with a 5 mm nozzle fueled with methane at Uo = 0.39 m/s, Re = 120, and coflowing air maintained at 0.15 m/s. When the fuel was modified by mixing it with an inert gas, the outer structure was suppressed and the flame bulge disappeared. This was observed when the 10 mm nozzle was fueled with a hydrogen and nitrogen mixture having a hydrogen to nitrogen volumetric ratio of 1 to 2.67, at Uo = 6.21 (Re = 3452, based on the mixture viscosity assuming an ideal solution). At this condition, the stoichiometric mixture fraction is 0.5282 (0.0283 for pure hydrogen diffusion flame). The suppression of the outer vortex (thus the flame bulge) is due to the change in stoichiometric condition that moves the flame location toward the inner boundary of the shear layer adjacent to the potential core. As a result, the near flame locations are no longer buoyancy dominated but momentum dominated; the outer vortex was suppressed. Consequently, the flame bulge was not observed. It was noted that further decrease (slightly) in the H_2/N_2 ratio results in a blow-off of the previously stabilized flame.

From the visualization and the spectrum analysis discussed above, it is clear that the flame induced buoyancy forces can exert significant effects on the flame dynamics. To appreciate the effects on the flame length (an obvious quantity of the flame), computation was made employing a modified GENMIX program similar to that used in.[23] The governing equations were cast in Favre averaged forms and model constants were adopted from.[24] To eliminate fuel density effects, ethylene diffusion flames were considered. The computation was made for a 10 mm nozzle, assigning a plug velocity profile and low turbulent intensity at the nozzle exit. A low velocity coflowing air was maintained at 8.3% of the

volumetric mean velocity of the fuel jet at the nozzle exit.

The results, Fig. 10, showed that the flame length, defined as the location of the mean stoichiometric condition, increased continuously as the fuel flow rate is increased when the buoyancy forces are included in the computation. This is, however, not the case when the gravitational acceleration was suppressed in the computation. As can be seen from the figure, the flame length first increased as the fuel flow rate was increased until transition to turbulence was encountered. Immediately beyond the transition, the flame length decreased as the fuel flow rate was increased. This is due to the enhanced transport resulting from flow turbulence. Further increase in the fuel flow rate does not result in a substantial decrease in the flame length. The no-buoyancy prediction is similar to what observed by Hottel and Hawthrone.[25] It is noted that the experiments reported in [25] involved small orifice nozzles and high fuel jet velocities. These flows, therefore, were momentum dominated with negligible buoyancy effects. In the present computation, due to the relatively large diameter was employed, the flame approached to the momentum dominated limit (for the visible flame region) at very high velocities (> 100 m/s). Figure 10 demonstrated the practical significance of the buoyancy effects in diffusion flames.

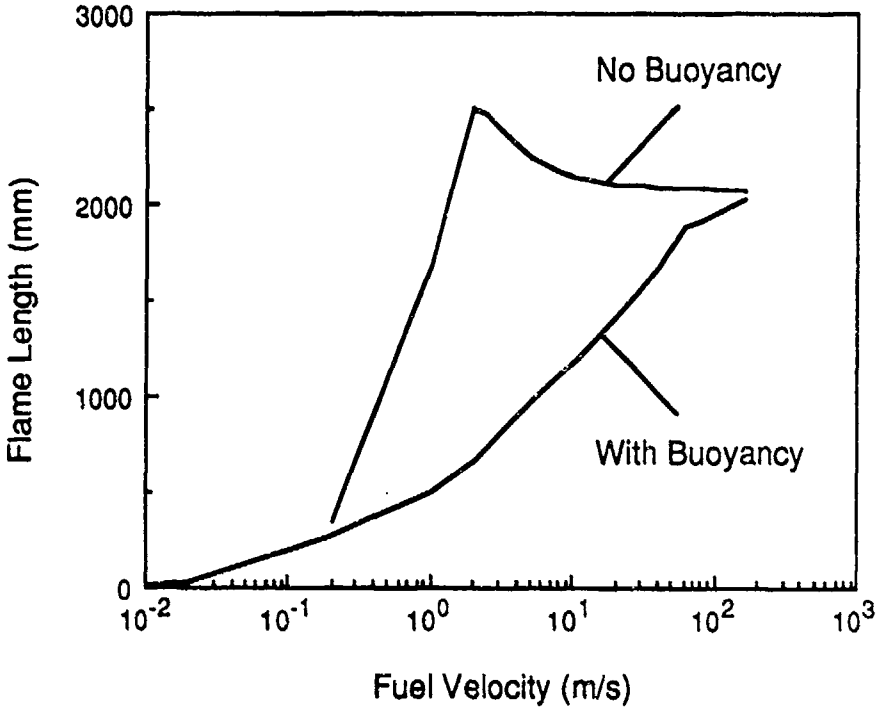


Figure 10. Flame Length of Ethylene Diffusion Flames (10 mm Nozzle)

SUMMARY AND CONCLUSIONS

In Refernce [10], Buckmaster and Peters showed that the flame oscillation is a result of buoyancy driven Kelvin-Helmholtz instability. In this study, the flow structure due to buoyancy driven Kelvin-Helmholtz instability and its relations with respect to the flame bulge are identified. The oscillation frequency was found in the range of 11.4 to 17.4 Hz and the convective velocity of buoyancy driven toroidal vortices was in the range of 0.66 to 0.87 m/s, over the conditions studied in this paper. The flame oscillation frequency showed a weak dependence on burner exit velocity and coflowing annulus air. The frequency was increased as the burner exit velocity or the coflowing air velocity was increased.

Buoyancy driven toroidal vortices are shown to exert significant stretching on the flame surface. The stretching can locally quench the otherwise continuous flame surface, resulting in a detached flame puff. The toroidal vortices are also responsible for the pairing or merging of the flame bulge. A sub-harmonic of the oscillation frequency was measured at locations where merging or pairing was observed. It is clear that buoyancy driven vortices are important to the dynamics of the flame, at least, at low and transitional Reynolds number conditions. To advance our understanding of the fundamentals of these flames, time resolved multipoint measurements will be required.

It should be mentioned that there exists a certain relationship between the flame oscillation and the vortex shedding inside the luminous flame for some diffusion flames. For example, vortex shedding, at a frequency of 13 Hz, inside a propane diffusion flame having Do = 22.5 mm, and Uo = 0.214 m/s (Re = 1100) was shown to be related with the flame oscillation.[21] The coflowing annulus air was maintained at 0.15 m/s in that experiment. Further study on the interaction between the buoyancy driven vortices and fluid motion inside the flame is needed to characterize the phenomenon.

The importance of the buoyancy driven Kelvin-Helmholtz instabilities to the structure of jet diffusion flames is recognized. To further our knowledge of the flame dynamics, detailed measurements of large diffusion flames are needed. As for the diagnostic tools, simultaneous multi-point measurements are in urgent need if one would like to advance the state of the knowledge on jet flames important to the Air Force propulsion applications.

Effort was also devoted to assemble the visualization data in a booklet form. A preliminary draft is enclosed in the Appendix. Color copies of the figures are available at AFWAL/POSF (Dr. W. M. Roquemore).

REFERENCES

1. Chamberlin, D. S. and Rose, A. (1928) 76th. Meeting of ACS, Swampscott, MA, 1928 (also in First Symposium on Combustion, The Combustion Institute, 1965, pp. 27-32).

2 Remenyi, K. (1980), Combustion Stability (Chapter 5, Flame Frequency), Akademiai Kiado, Budapest, pp. 60-96.

3. Kimura, I.(1965) 10th Symposium (international) on Combustion, The Combustion Institute, pp. 1295-1300.

4. Grant, A. J. and Jones, J. M. (1975) Combust. Flame 25 153-160.

5. Ballantyne, A. and Bray, K. N. C. (1977) 16th Symposium (international) on Combustion, The Combustion Institute, pp. 777-787.

6. Yule, A. J., Chigier, N. A., Ralph, S., Boulderstone, R., Ventura, J.(1981) AIAA J. 19 752-760.

7. Becker, H. A. and Liang, D. (1983) Combust. Flame 52 247-256.

8. Eickhoff, H. (1982) in Recent Contributions to Fluid Mechanics (W. Hasse, ed.), Springer-Verlag, pp. 50-57.

9. Durao, D. F. G. and Whitelaw, J. H. (1974) Pro. Royal Soc. London A 338 379-501.

10. Buckmaster, J. D. and Peters, N. "Infinite Candle and Its Stability-a Paradigm for Flickering Diffusion Flames", in 21st Symposium (international) on Combustion, The combustion Institute (in press).

11. Eickhoff H., and Winandy, A. (1985) Combustion and Flame, 60, 99-101.

12. Strawa, A. and Cantawell, B. J. (1985) Phys. Fluids 28 2317-2320.

13. Roquemore, W. M., Cher., L.-D., Seaba, J. P., Tschen, P. S., Goss, L. P., Trump, D. D. (1987) Physics Fluids, Vol. 30 p. 2600.

14 Roquemore, W. M., Chen, L.-D., Goss, L. P, and Lynn, W. F. (1987) "Structure of Jet Diffusion Flames," Proceedings of Joint U.S.-France Workshop on Turbulent Reacting Flows, Rouen, France, 6-10 July 1987 (Also to Appear in the Lecture Notes in Physics Series, Springer-Verlag, 1988).

15. Roquemore, W. M., Tankin, R. S., Chui, H. H., and Lottes, S. A.(1986) Exp. Fluids 4 205-213.

16 Chen, L.-D. and Roquemore, W. M.(1986). Combust. Flame 66 81-87

17. Barin, I. and Hnacke, O. (1973) Thermochemical Properties of Inorganic Substances, Springer-Verlag, p. 762.

18 Gordon, S. and McBride, B. J. (1976) "Computer Program for Calculation of Complex Chemical Equilibrium," NASA SP-273, NASA.

19. Freymuth, P., Bank, W., and Palmer, M. (1985) in Flow Visualization III (W. J. Yang, Ed.), pp.

99-105.

20. Durst, F., Melling, A., and Whitelaw, J. H. (1976), Principles and Practice of Laser-Doppler Anemometry, Academic Press, pp. 272-294.

21 Chen, L.-D. and Roquemore, W. M. (1987), Proceedings of ICALEO 1986 Vol.: 58 Flow and Particle Diagnostics, (Fifth International Congress of Applications in Lasers and Optics, 10-13 November 1986), pp. 16-23, Laser Institute of America, 1987.

22. Perry, A. E. and Chong, M. S. (1987), Ann. Rev. Fluid Mech., Vol. 19, pp. 125-155.

23. Jeng, S-M., Chen, L.-D., and Faeth, G. M. (1982), Nineteenth Symposium (International) on Combustion, The Combustion Institute, pp. 349-358, 1982.

24. Jeng, S-M. and Faeth, G. M. (1984), ASME J. Heat Transfer, Vol. 106, pp. 721-727.

25. Hottel, H. C. and Hawthorne, W. R. (1949), Third Symposium (International) on Combustion, The Combustion Institute, pp. 255-266.

Appendices can be obtained from

Universal Energy Systems, Inc.

FINAL REPORT NUMBER 9
REQUESTED A NO-COST TIME EXTENTSION
TO BE SUBMITTED IN 1987 MINI-GRANT FINAL REPORT
Dr. Hoffman Chen
760-6MG-118

FINAL REPORT NUMBER 10
REPORT NOT RECEIVED IN TIME
WILL BE PROVIDED WHEN AVAILABLE
Dr. John Cheung
760-6MG-135

1986 USAF/UES MINI GRANT PROGRAM

Sponsored by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the
UNIVERSAL ENERGY SYSTEMS, INC.

FINAL REPORT

INFRARED FLUORESCENCE AND PHOTOACOUSTIC MEASUREMENTS OF NO(v=2) RELAXATION AS A FUNCTION OF TEMPERATURE

Prepared by:	Dr. Wolfgang Christian
Academic Rank:	Associate Professor
Department:	Department of Physics
University:	Davidson College
Research Location:	Davidson College Davidson N.C. 28036

Date:	December 31, 1987
Contract Number:	F49620-85-C-0013/SB5851-0360
Purchase Order Number:	S-760-6MG-030

Abstract:

A high power infrared source using Stimulated Electronic Raman Scattering (SERS) in heat pipe ovens was developed in order to excite rotational transitions in gases of atmospheric interest. The source was able to produce intense IR radiation in the 2.8, 7.2, and 13.0 um regions of the spectrum but the tuning range was considerably less than expected. Consequently, only a very limited number of experiments were possible with N_2O and the proposed NO experiments could not be carried out at all. A survey of IR emission and multiphoton ionization in Cs was carried out in an attempt to find other suitable IR producing mechanisms for vibrational excitation of NO.

Six physics majors taking Intermediate Laboratory were trained to operate the laser system and used it to measure the quantum defect, SERS IR generation near the 7P,8P and 9P levels, and other nonlinear effects in Cs. Two of these students were hired back during the summer to work as laboratory assistants under the Air Force Office of Scientific Research Mini Grant in order to complete the study of Multiphoton Ionization and IR generation in heat pipe ovens.

Summary of Completed Project

Laser-induced ionization and Stimulated Electronic Raman Scattering (SERS) was studied in Cs vapor from room temperature (~10^{-6} Torr) up to 500 K (~10 Torr) for cases in which the laser frequency is tuned at or near the $np^2P_{3/2,1/2}$ (n=7,8,9) levels in an attempt to produce a usable high power IR source for the excitation of molecules of atmospheric interest. At low atom number densities ionization occurs as a result of resonantly enhanced Multi-Photon Ionization (MPI). At high cesium densities (1 - 10 Torr) and laser powers, SERS emission is observed originating from near the $7p^2P$ and $8p^2P$ levels. Ionization occurs over a broad range of laser wavelengths and follows the SERS signal including a dip at line center as shown in figure 1. The ionization may be caused by chemi-ionizing collisions between two fast, excited $6p^2P_{3/2,1/2}$ atoms produced in the SERS process in addition to MPI. A major contribution of the dip is due to an interference between the one-photon and a three-photon coherent process. Competition between SERS, MPI and other processes may in fact limit the conversion efficiency and tuning range of SERS and

make it difficult to produce the 2.7 um radiation necessary to excite the v=2 vibrational state in NO. The first section presents limited results obtained with vibrational excitation of N_2O using the SERS source; a second section presents a survey of IR emission and MPI in Cs heat pipe ovens that was conducted to look for other IR generating mechanisms.

SERS Experiments

IR Emission and Tuning Range:

The IR tuning range of the emitted SERS radiation was measured near the 7P, 8P, and 9P levels in Cs and 2.7 um, 7.2 um and 13 um radiation was produced. Figure 2 shows the energy level diagram for Cs and the pumping mechanism for SERS while figure 3 shows the IR emission obtained near the 7P level by placing a HgCdTe detector (shielded from visible light with a Ge flat) directly behind the heat pipe oven. The apparatus shown in figure 4 was similar to the one constructed at AFGL in 1986. The higher powers of our Nd:YAG pumped dye laser extended the IR tuning range to past 4.2 um on the red side of the 7P intermediate state but produced no enhancement on the blue side from that reported by Cotter and Hanna. Spectrographic analysis, figure 5 and 6, of the IR beam using an SA Instruments 1/2 m monochrometer showed that the IR beam is spectrally dirty due to SERS from the OPSE emission present in the beam. Consequently, it is only possible to use SERS as an IR source if the laser is very close to resonance with an intermediate P state. Although it was not possible to study the NO (v=2) transition, a study was performed on N_2O in order to show the feasibility of IR spectrographic analysis using an SERS source.

N_2O Measurements

The (1,0,1) and (0,2,1) bands in N_2O were excited and the IR transmission and photoacoustic (PA) spectrum were measured. Figure 7 shows the incident IR beam and clearly demonstrates the

large variation in Raman gain near the 7P states. The transmitted IR beam shown in figure 8 is even more complex since the IR absorption also varies rapidly in this region. The photoacoustic spectrum shown in figure 9, however, is remarkably clean since the microphone only responds to absorbed energy. The (1,0,1) and (0,2,1) transitions are clearly visible and although the shape is suspect due to the rapid variation of the incident IR beam, SERS can be used for kinetic studies of these states.

The Nd:YAG pumped dye laser was tuned to produce strong SERS IR emission in order to excite the P branch of the (0,0,0)-->(1,0,1) vibrational transition in N_2O. Following this excitation, a time resolved photoacoustic signal was obtained using a Tektronix 7D20 signal averager. Since the PA signal is directly proportional to pressure, the time resolved signal can be used to determine the rate at which energy is deposited in the translational degrees of freedom via V-->T relaxation. Unfortunately, microphone and cell effects modify the signal. Figures 10 and 11 show spectra at 2 and 5 Torr respectively. Although the increase in the collisional relaxation time is clearly observable, ringing in the cell of transverse acoustic resonances and the more complicated kinetics of N_2O as opposed to NO make it impossible to obtain useful molecular information from the data.

MPI and IR Emission in Heat Pipe Ovens

Phenomena other than simple SERS such as 4 wave mixing, hyper-SERS, or excimer emission may produce strong IR emission and warrant further study. Consequently, a survey of MPI and IR emission in a Cs heat pipe oven was conducted. The results of this survey are presented below.

DCM:

The DCM experiments begun at AFGL in 1986 were continued at Davidson College. The addition of a UV WEX allowed excitation

near the one photon ionization limit while the addition of a 1/2 m IR spectrophotometer allowed spectrographic analysis. The DCM three photon ionization signal, figure 12, showed enhancement at the two photon D and S dipole resonances while the doubled DCM MPI signal, figure 13, showed enhancements at the one photon dipole transition to the P states and quadrupole transition to the D state. An unknown peak at 6594 A has a shape characteristic of excimer emission but no IR excimer emission was found in the 1.5 um to 15 um range. The IR emission shown in figure 14 was due primarily to fluorescence from excited Rydberg states in Cs; the strong IR emission observed between Rydberg states at the AFGL data could not be duplicated.

Rhodamine 640:

The R640 MPI scan, figure 15, shows a hybrid Rydberg series originating on the $5D_{3/2}$ and $5D_{5/2}$ excited state. Both Raman scattering to and photoionization of the Cs dimer could produce these hybrid resonance. The absence of tunable IR emission near 5 um and the presence of a strong dimer absorption band at 6300 A are strong indicators that photo-ionization in the most likely mechanism for production of the 5D excited states.

LDS 698:

LDS 698 signal shown in figure 16 has a whole range of one photon, two photon and hybrid resonances. Figure 17 shows that many of these resonances have strong IR emissions produced by laser action or fluorescence from excited Rydberg states. The doubled LDS 698 scan, figure 18, shows a very wide MPI peak near the 9P one photon resonance due to the SERS into the 9S state followed by one photon ionization. This process is identical to the the Coumarin 460 scan near the 7P states and produced 13 um raditation.

Student Research:

Eight physics majors were trained in laser spectroscopy as part of the AFOSR-UES mini grant. Junior laboratory students were

broken up into three teams and assigned two week projects in rotation. The first group used a heat pipe oven to measure the SERS gain in a heat pipe oven when the dye laser was tuned near the $7P_{3/2,1/2}$ state in cesium. The second group measured the energies of the high S and D (n = 13, 14, ...30) Rydberg states using two photon resonant three photon ionization. The third group measured four wave mixing between the IR beam generated through SERS and two pump dye laser photons. These results were then presented as poster sessions during the final week of classes as part of a departmental symposia.

More advanced work was done by Mrs. Barbara Brooks who completed a survey of ionization effects in the visible and near IR using two and three photon ionization of cesium. During the summer of 1987 two students, Mr. Sim Harbert and Mr. Eddie Alexander were hired as research interns to help study the feasibility of using SERS as a high power IR source for spectroscopic experiments.

Papers Presented:

W. Christian, "Multiphoton Ionization and Infrared Generation In A Cesium Heat Pipe Oven Near The Two Photon Ionization Limit" International Conference on Multiphoton Processes IV, Boulder, CO, July 13-17, 1987.

J. A. D. Stockdale, R. N. Compton, Adila Dodhy, W. Christian, P. Lambropoulos and T. Olson, "Laser-Induced Ionization and Stimulated Electronic Raman Scattering In Cesium Vapor Near The $np^2P_{3/2,1/2}$ (n=6,7,8,9) States", International Conference on Multiphoton Processes IV, Boulder, CO, July 13-17, 1987.

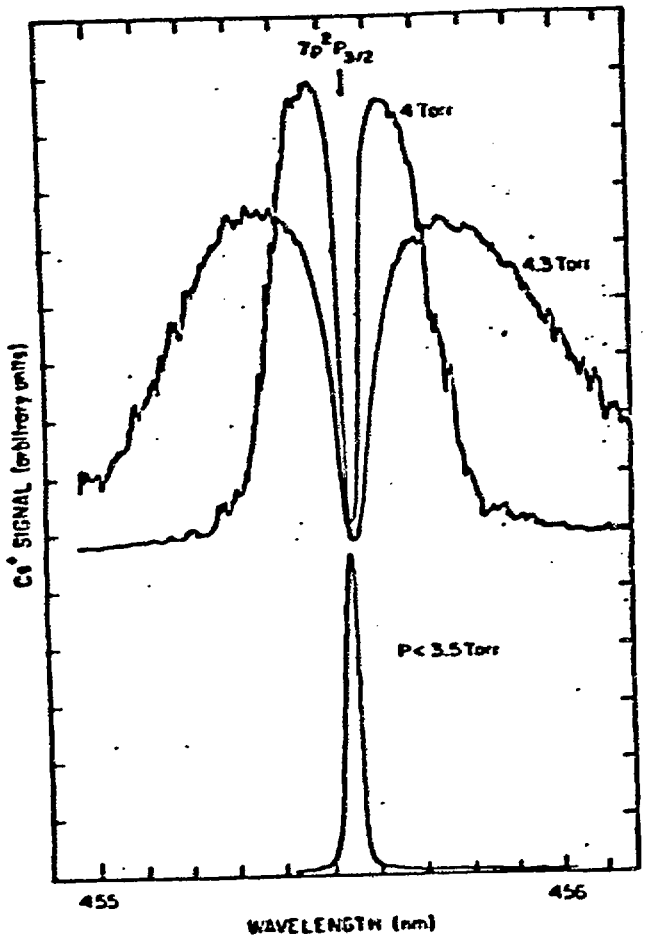


Figure 1. Ionization spectrum in the vicinity of the one-photon allowed $7p^2P3/2$ state of cesium at high and low atom number densities.

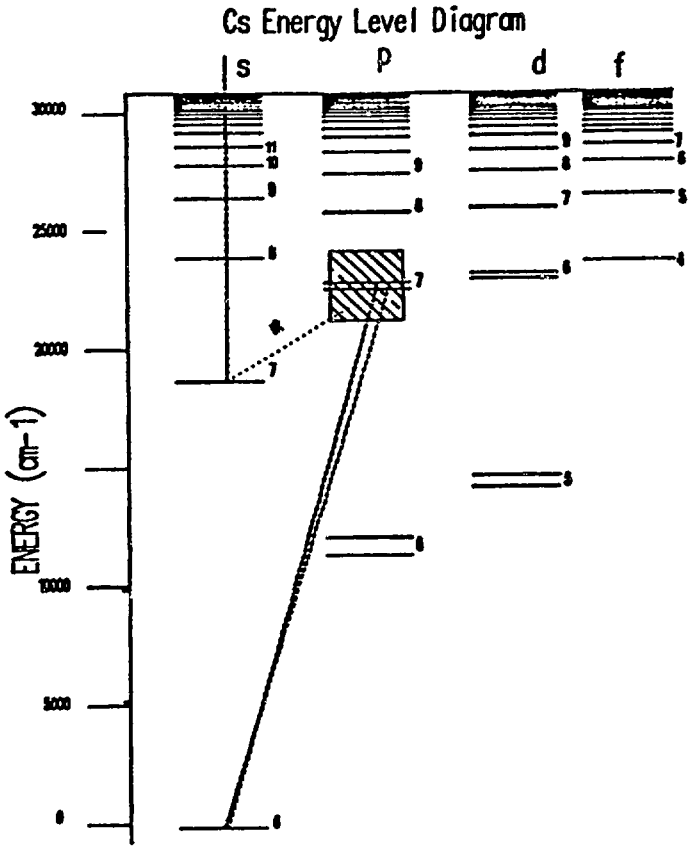


Figure 2. Cesium energy level diagram showing pump mechanism for SERS and subsequent photoionization from the 7 S state.

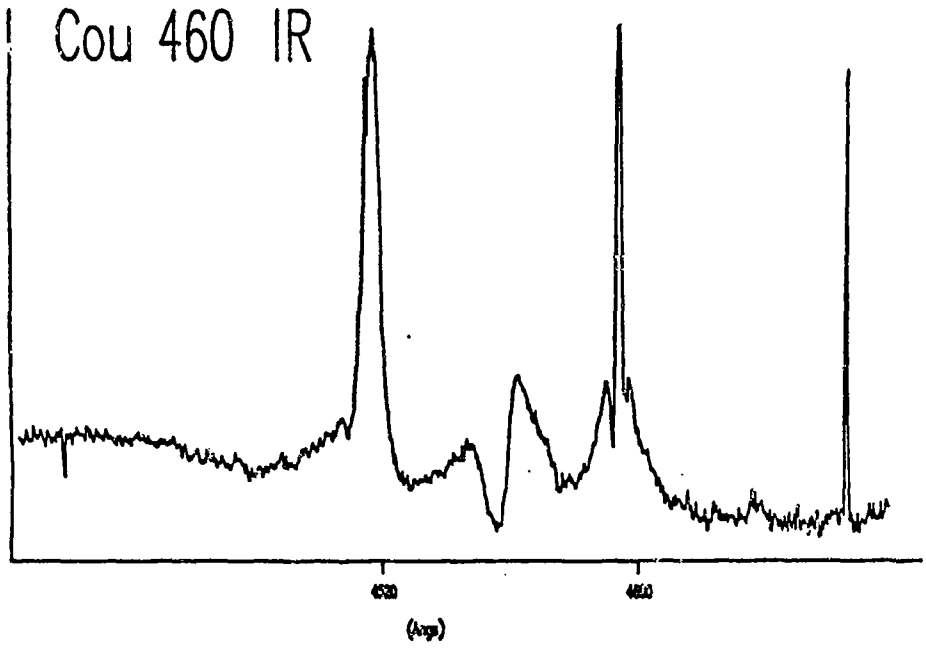


Figure 3. Total IR emission from heat pipe oven due to SERS and other processes.

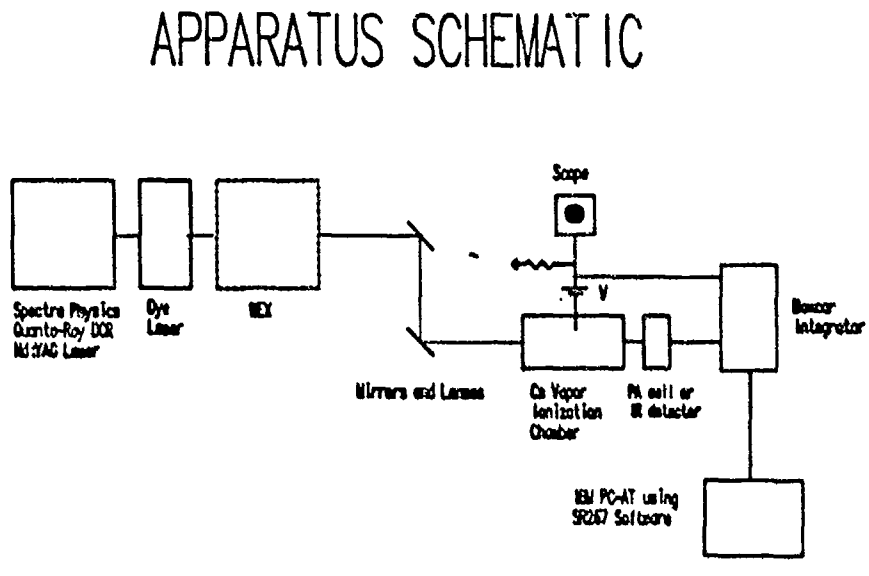
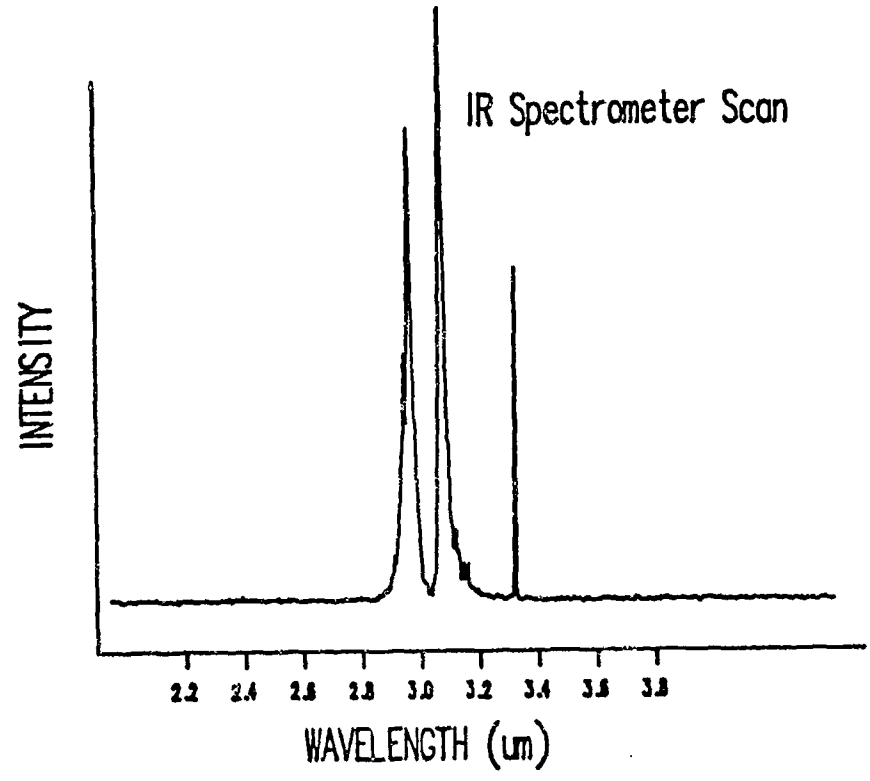
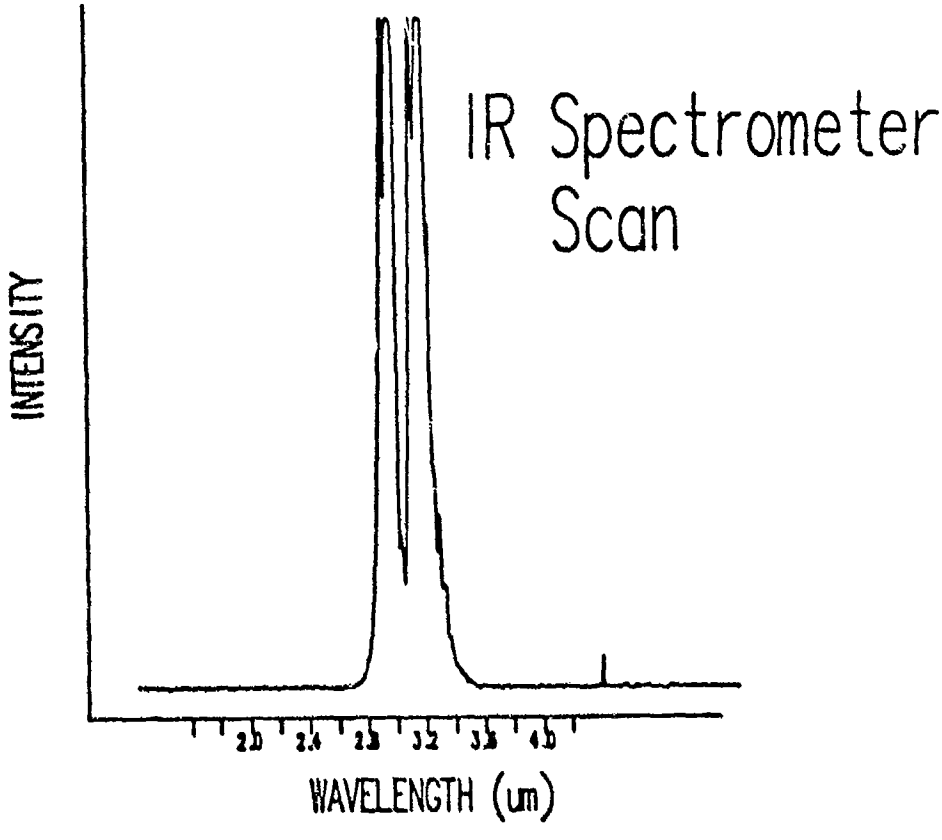


Figure 4. Experimental setup used for MPI and IR generation experiments.





Figures 5 and 6. IR spectrometer scan with pump laser tuned far from 7P transition. The small peak to the right in each scan is SERS from the dye laser beam. The broad peak is SERS from the OPSE of the dye.

Figure 7

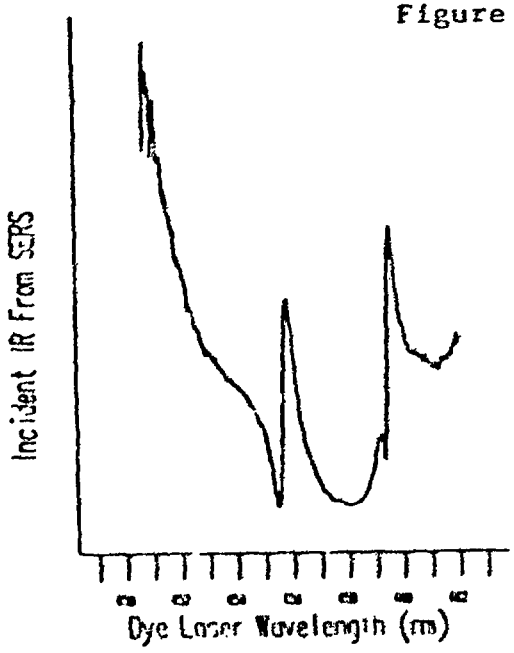


Figure 8

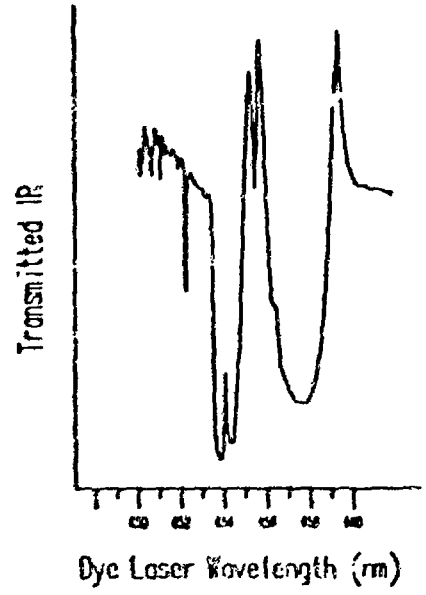
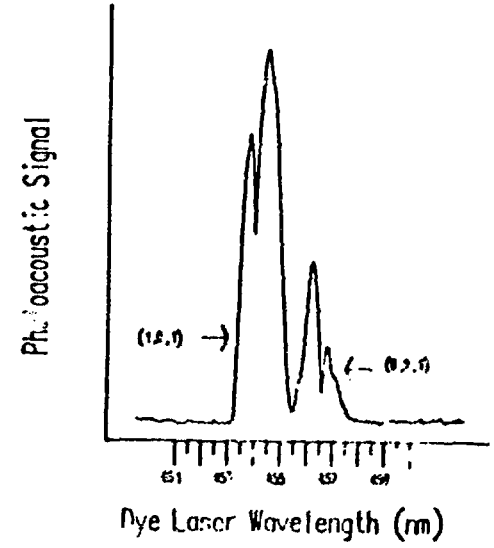


Figure 9



Figures 7, 8, and 9. The IR produced by the heat pipe and incident on the cell, the IR transmitted through the N_2O cell, and the photoacoustic signal inside the cell.

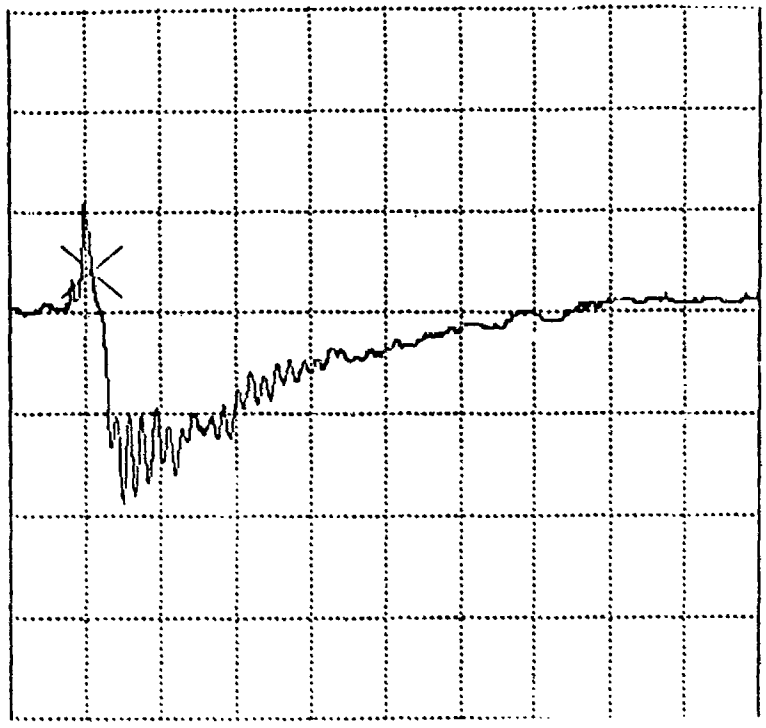


Figure 10

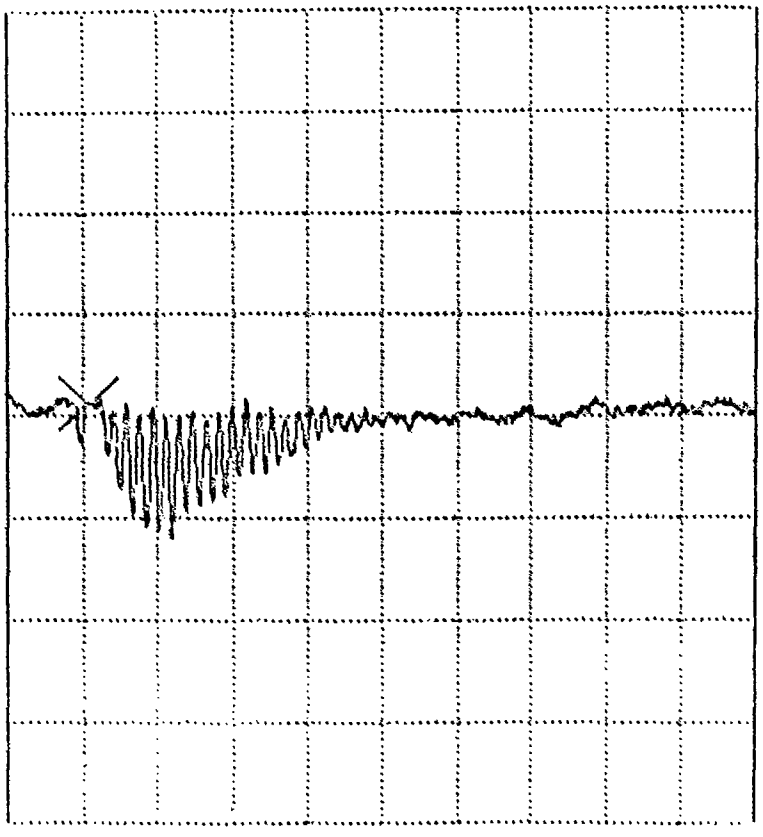
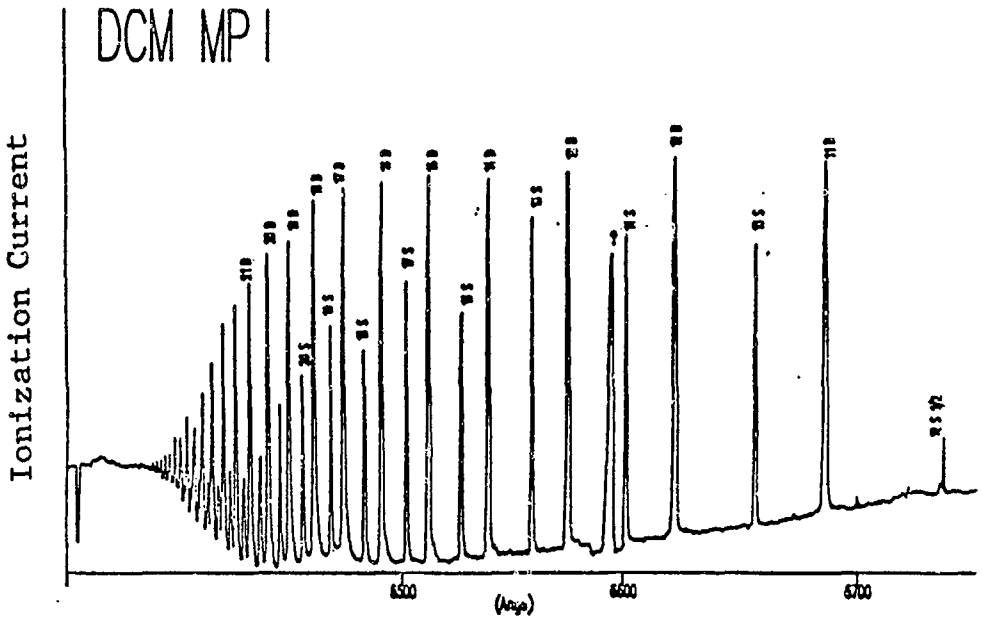
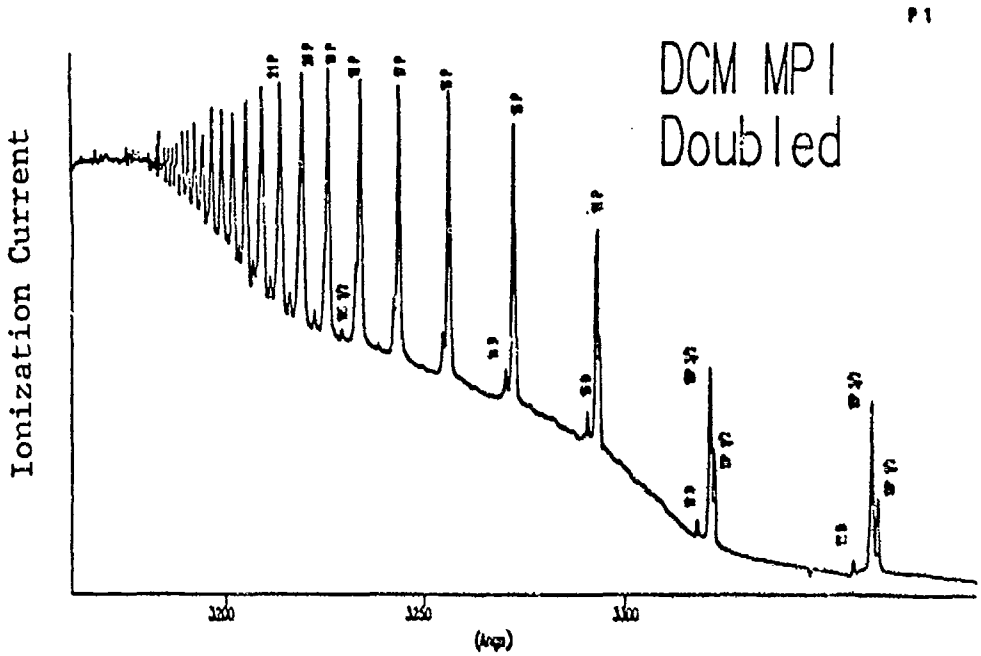


Figure 11

Figures 10 and 11. The time resolved photoacoustic signals at 5 Torr and 2 Torr, respectively. The high frequency component of the signal is due to ringing of the fundamental transverse acoustic mode of the cell. Horizontal scale - 0.512 msec/div.





Figures 12 and 13. MPI in Cs heat pipe oven at 3.00 Torr as a function of incident laser wavelength.

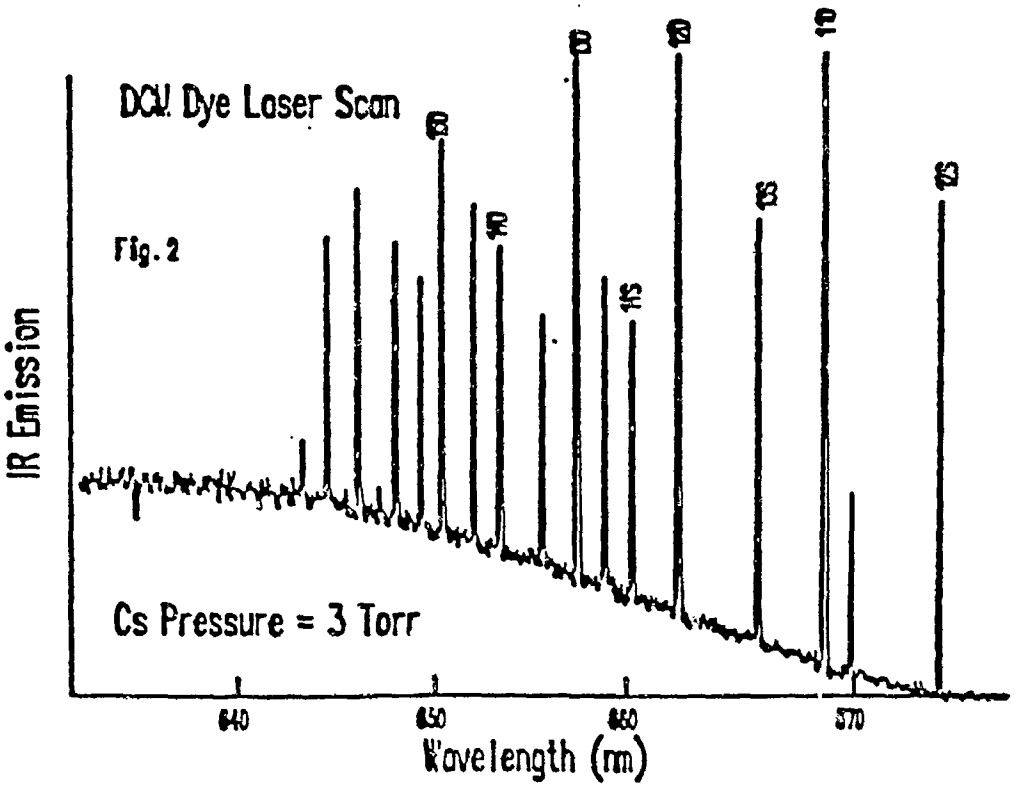


Figure 14. IR emission following two photon excitation of high Rydberg states in Cs heat pipe oven.

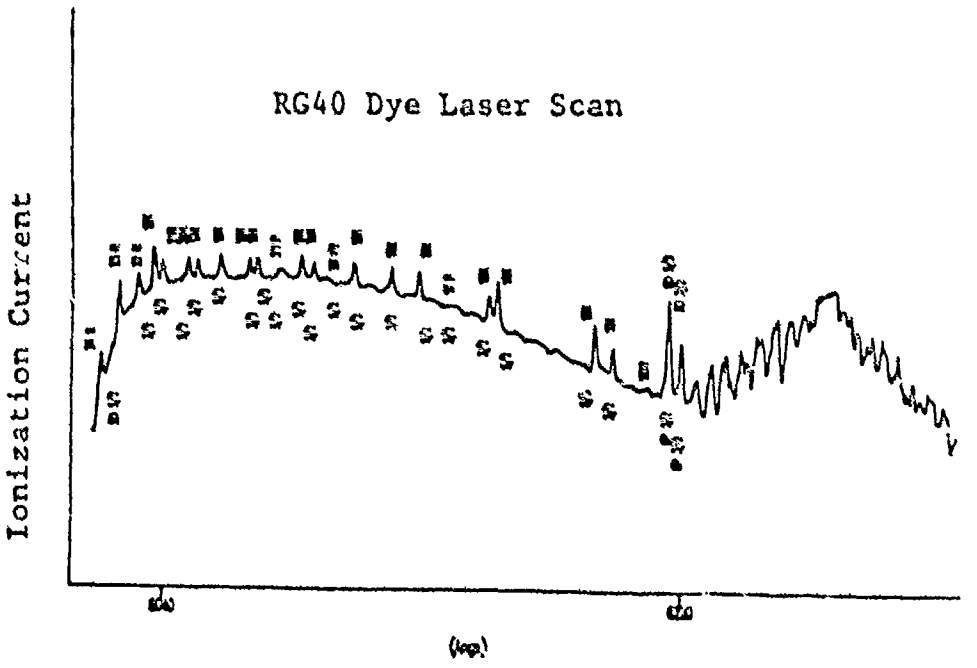
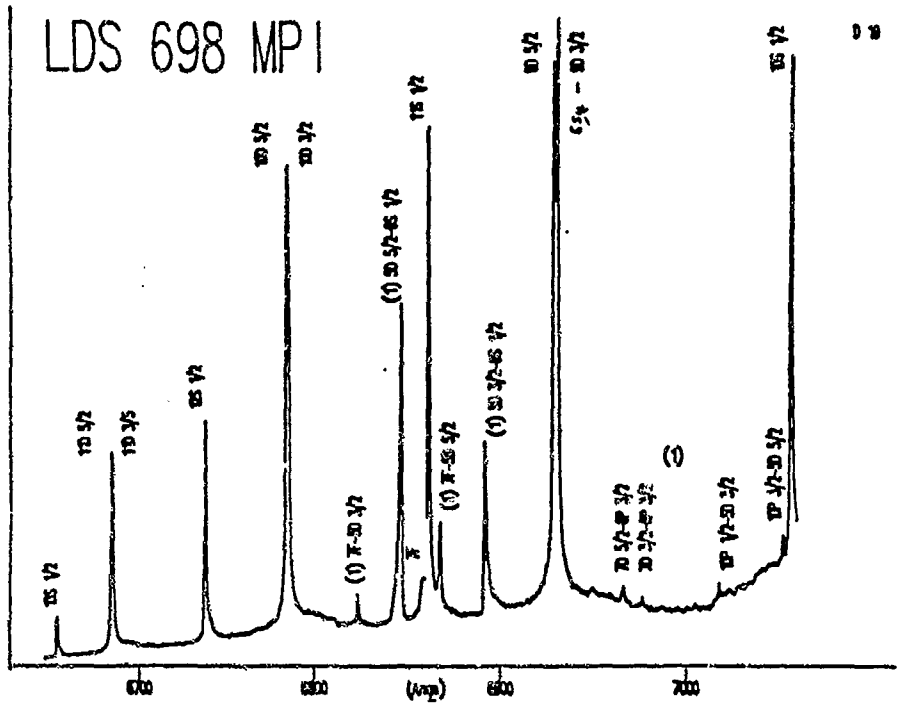
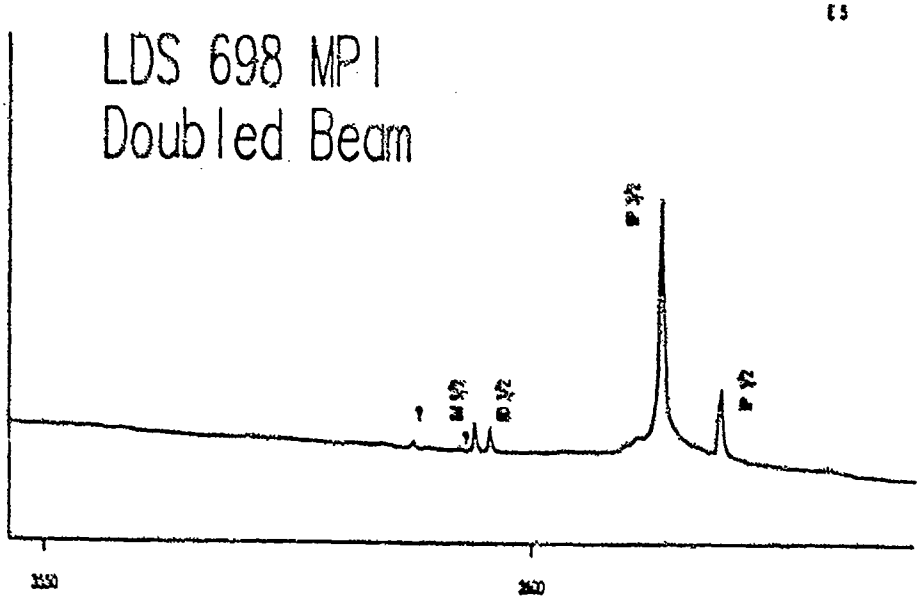


Figure 15, Cs MPI showing hybrid Rydberg series originating on the $5d^2D3/2$, 5/2 excited states.





Figures 16 and 17. Cs MPI in heat pipe oven as at 3.00 Torr as a function of incident dye laser wavelength.

USAF-UES FACULTY RESEARCH INITIATION PROGRAM

sponsored by

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

conducted by the
Universal Energy System, Inc.

FINAL REPORT

Heat and Mass Transfer in a Dual-Latent Heat Packed Bed Thermal Storage System

prepared by:	JACOB N. CHUNG (Principal Investigator) MARK M. WEISLOGEL (Graduate Student)
Department & University:	Dept. of Mechanical and Materials Eng., Washington State University
USAF Lab:	Thermal Energy Lab., Nuclear/Thermal Sys., Aerospace Power Division, Aero Propulsion Lab., Wright-Patterson AFB, OH 45433
USAF Researcher:	Mr. John E. Leland
Date:	February 5, 1988
Subcontract No:	S-760-6MG-067

(Video "Summary and Visualization" included with report.)

ABSTRACT

Experimental investigations on the condensation heat and mass transfer between flowing steam and a single encapsulated sphere filled with PCM (phase change material) and between flowing steam and a packed bed of the same spheres, were performed. The objectives of the research were to obtain transient transport characteristics of the dual latent heat thermal storage system. A special device was used to instrument the inside of the sphere to simulate the gravity-free melting environment for the PCM as the bulk of the heat transfer resistance was identified to be with the PCM. This device prevents the sinking of the unmelted solid PCM toward the bottom of the spherical shell, thus also eliminating the strong dependence of the flux on the tangential position. This therefore enabled the single heat flux transducer readings to represent closely the average heat flux for the entire sphere.

Instantaneous heat transfer rates were obtained for spheres of three different sizes with single sphere arrangements and a packed bed environment. The special device worked well for the larger spheres ($2.5''$ and $3.0''$), that allowed the normalized thermal energy stored to be correlated by a single dimensionless time scale.

Acknowledgments

We would like to thank the Air Force Systems Command and the Air Force of Scientific Research for sponsorship of this research. Additional support and guidance from the Thermal Energy Laboratories at Wright-Patterson AFB is also appreciated. Cost sharing, equipment, and computing time provided by the Department of Mechanical and Materials Engineering and College of Engineering at Washington State University have been essential for the completion of the project. Contributions from many M.E. undergraduates are also noted.

1. INTRODUCTION

Energy management systems are being used by industrial commercial establishments and military operations to reduce their overall energy consumption. Quite often many different sources and demands for energy exist and the total energy usage of the system can be reduced by the utilization of the waste heat from one or more processes as an energy source for others. Since the availability and the demand for the thermal energy may not coincide time-wise, the flexibility of the energy utilization system can be greatly increased by the development of an efficient method for storing thermal energy to insure the continuity of the thermal process.

Thermal energy storage can also be used to reduce the mass of the heat rejection system in pulled space power supplies. Waste heat generated during high-power sprint mode operation is reflected to the thermal storage and then alter dissipated to space via the radiator over the much longer non-operation period of the orbit. Thus, a significantly smaller capacity radiator is required.

There are many methods available for the storage of thermal energy. Some methods require the transformation of the thermal energy into other energy forms for storage purpose while others store the thermal energy directly. Examples of the first class of storage units include chemical systems, flywheels, batteries, compressed air, superconducting magnets, and underground pumped hydro systems. In direct thermal energy storage units the sensible heat of a liquid or a solid, the latent of fusion or evaporation, reversible chemical heat absorption, heat of hydration, or heat of chemical change are used.

Also, there are various forms of the storage phase design. Among them, the packed bed has been gaining most of the attention due to its high efficiency as a storage medium.

A pack bed system has also been used widely in many chemical engineering processes for the purpose of heat and mass transfer. The advantages of a packed bed comes mainly from its large interfacial transport areas during the processing. In such a system, a fluid is either heated or cooled by thermal convection as it passes through a mass of packed material with a certain porosity.

In a space-based thermal system, a thermal storage unit with high-energy storage density over a narrow temperature range is highly desirable. Under such requirements, a packed bed of spheres filled with a phase-change material (PCM) coupled with a phase-changed liquid metal working fluid is thought to be a best qualified candidate.

The unique features of fluid flow and heat transfer between the working fluid and the packed bed are listed as follows:

(i) Because of the isothermal melting process inside the encapsulated spheres there will be a section of the bed with constant temperature and the size and location of this "mushy" region vary as a function of time.

(ii) On the working fluid side, there also will be a 2-phase zone of vapor and liquid mixture. This zone will be moving along the flow direction in the packed bed.

(iii) The condensation of the working fluid on the spheres will take place as long as the sphere surface temperature is below the saturation temperature corresponding to the local pressure. Condensate will be trapped between the spheres. This will inhibit flow of the working fluid and reduce condensation rates.

Both experimental and analytical investigations become more difficult and complicated as compared with non-phase change systems. Comprehensive investigations which explore the above features and provide detailed results are currently unavailable in the published literature. This report documents the results of the experimental project which

is aimed at the understanding of the basic heat and mass transfer mechanisms in a dual-latent heat packed bed thermal storage system.

The following is a list of the specific results obtained in the research work.

1. The instantaneous heat transfer rates between the condensing fluid and the PCM inside the sphere.
2. The effects of the size of the encapsulated sphere on the heat and mass transfer rates.
3. The differences in transport rates between a single sphere and a typical sphere in a packed bed.

2. LITERATURE REVIEW

In the literature, a considerable amount of work is concerned with the heat and mass transfer in packed beds used as chemical reactors. Those reports mainly focus on the steady operation of the process. In this section, we restrict our view only on the thermal storage systems using the packed bed as the storage medium. Most of the investigations have been analytical and numerical in this area. Local equilibrium models assume that the instantaneous solid and fluid temperatures to be equal. Two-component models allow these temperatures to be different, with the interphase heat transfer described by a mean heat transfer coefficient. The majority of work has focused on the one dimensional, time-dependent, two-component model first developed by Schumann (1928) for fairly simple initial and boundary conditions. This model predicts the mean solid and fluid temperatures as a function of axial location and time. Shitzer and Levy (1983) extended the basic Schumann analysis to account for a time variation in inlet temperature using the Duhamel superposition technique. Their results, which compare favorably with

experimental measurements, are applicable to a charging mode only. Some analyses have attempted to account for two- and three-dimensional effects. Martin (1978) divided a cylindrical bed, packed with uniform spheres, into a central core and a near-wall region, with a higher void fraction near the boundary than in the core of the bed. He used an assumed radial variation of void fraction to predict a cross-sectional averaged heat transfer coefficient. Gross, et al. (1980) utilized Martin's correlation to predict the axial variation of temperature within the bed. Their explicit numerical method in most applicable to liquid systems.

Amundson (1956) analyzed the two-dimensional temperature distribution in a cylindrical packed bed chemical reactor with energy losses at the wall, and energy generation in the solid phase. For constant velocity, void fraction, fluid properties, and transport coefficients, the transient solutions for various boundary and initial conditions are obtained. The effect of including an axial diffusion term is examined. The results are mathematical in form without experimental verification. Robertson and Cavendish (1981) examined the two-dimensional response of catalytic converters and packed beds for the case of uniform void fraction, but non-uniform inlet conditions. Recently, Beasely and Clark (1984) used a finite difference method to study the transient response of a packed bed thermal storage unit. They specially emphasized on the effects of axial and radial dispersions with arbitrary time and radial variations in the inlet fluid temperatures.

On the experimental investigations, a series of experiments was conducted by Peavy and Dressler (1977) who considered transpiration heat transfer of water in a sand column. Jones and Hill (1979) performed experiments in a pebble-bed composed of 3.8 to 5 cm diameter washed river gravel packed into a well-insulated bed with a square cross-section

measuring 1.98 m on a side and 1.83 m in length. The packed bed was subjected to a step change in inlet conditions by passing air. They measured the fluid temperatures and the bed temperature responses. Beasely and Clark (1984) also did an experimental study to compare the results with their own numerical simulations. They measured temperature distributions in both axial and radial directions to investigate the responses of the bed to variations in inlet air temperatures.

So far we have reviewed literatures that deal with systems with no phase-change of either the working fluid or the bed material. Yimer, Crisp, and Mahefkey (1980) and Kordahi, Yimer, and Crisp (1982) have performed numerical simulations of the melting process inside an encapsulating cylindrical material for a thermal storage system. For melting in a spherical enclosure, Moore and Bayazitoglu (1982) developed a numerical model to predict the heat transfer rates for the case of a constant enclosure surface temperature. They reported the time dependent interface position and the energy stored as a function of time. A good comparison was found between the numerical predictions and the experimental results.

Torab and Beasely (1985) employed a finite difference method to study the dynamic response of a packed bed of encapsulated PCM. They found the bed with the PCM stores 2.85 times more energy but requires a longer charging period.

Chung (1986) used finite difference numerical model to study the fluid flow and heat and mass in a dual-latent heat packed bed thermal storage system. In his numerical model, the encapsulated spheres are considered as volumetric heat sinks for the condensing vapor working fluid. Because of the lack of information on the overall heat transfer coefficient between the working fluid and the encapsulated spheres, a parametric study was performed

to investigate the dynamic response of a dual-latent heat packed bed thermal storage system consisted of with 1″ diameter spheres filled with LiH as the PCM and Nak as the working fluid. High Reynolds number flows result in shorter time required for fully charging the bed. In the open literature, there is no single study which deals with the experimental investigation of the phase-change of the working fluid in a packed bed.

It is believed that condensation of vapor inside a packed bed is a complicated phenomenon. Careful measurements and observations are needed to understand the transport mechanisms of heat, mass, and momentum. Experimentation of vapor condensation inside a packed bed with PCM offers one definite advantage that there will be a substantial period that vapor condensation will take place on isothermal spheres. This isothermal period will provide a stable period for accurate measurements of condensation rates of the vapor flow.

3. EXPERIMENTAL APPARATUS

A general schematic of the experimental apparatus employed in this study is provided as Figure 1. The primary thermal loop, capable of both opened and closed cycle operation, consists of a metering flow pump, two boilers in series, a dummy chamber to establish the working fluid condition, a chamber to house the spheres under study, and a condenser. Photos of each individual component mentioned above are included (see attachments), and are numbered accordingly as presented below. Their respective functions will be addressed in the procedures section.

3.1 System Pump.

The pump selected for our study was a Masterflex metering pump manufactured by

Cole-Parmer Inc. The advantage of this pump was its ability to deliver a wide range of steady flows regardless of the pressure conditions up or down stream. The mass flow rate delivered by the pump could be accurately determined by multiplying the rotational speed of its drive shaft by a constant depending on the size of tubing used. The pump was self priming and had the added benefit of providing no opportunity for leakage of air into the system.

3.2 Heat Input.

The pump circulated condensed liquid to a 3500 Watt preheater and then on to a 3800 Watt boiler. Both heat sources were variable coil-types, the former being controlled via a temperature feedback loop and the latter by a rheostat. Though only one inlet steam condition was used for this study, with the flow rate and heat input so variable, a much wider range of data may be obtained in the near future. Exit steam temperature from the second heater was monitored by a thermocouple which was vital during start-up and shut-down of the system. The heaters were well insulated during each of the tests.

3.3 Chambers and Sphere Holders.

Two chambers were implemented whose purposes will be discussed shortly. The aluminum chambers were identical in geometry except for a few access ports on one. O-ring seals allowed for moderately high vacuum conditions. Both chambers were leak tested to less than $10\mu m$ Hg. A more detailed schematic of this part of the cycle is provided as Figure 2 with rough dimensions given. Wire screens in the entrance cones of each chamber acted to shed to the perimeter of the main chamber bodies any condensate which might have otherwise come in contact with the spheres or ambient temperature and pressure probes. A similar screen located in the exit cones of each chamber was used as a "safety

net" in the case of a sphere mishap.

As pictured in the figure, aluminum sphere holders were designed and fabricated which could be inserted into the chambers. Each of these holders were intended to support the single sphere and multiple sphere arrangements and were sized accordingly. The spheres of the arrayed tests were supported on the sides by the aluminum insert and on the bottom by steel rods which were isolated thermally by wrapping them in 1/8" Tygon tubing. The chambers and sphere holders were made of 1/8" and 1/16" thick sheet respectively. Both chamber were insulated with thermal blankets to minimize wall condensation and heat loss to the surroundings.

3.4 Condenser.

A simple coil condenser was employed for closed loop operation of the system to guarantee liquid return to the pump. Estimated to be able to condense 10 KWatts of steam vapor, this item was simply bypassed for most of the experimentation since the open cycle was found more suitable for our purposes.

3.5 Vacuum System.

The entire apparatus could be evacuated by a vacuum line connected at one of the equipment ports of the "Real" chamber. A 1 Hp two stage vacuum pump provided the suction for the experiments and could bring the system of chambers, 3/8" copper tubing, brass Swage-lock fittings, and quarter turn globe valves to 380μm Hg (roughly 0.05% of an atmosphere). Higher vacuum was obtainable if confined to the chambers (roughly 100μm Hg). This phenomena was exploited as will be described shortly.

Low vacuum pressures were monitored by thermocouple gauges whose analog output was read from a Veeco vacuum chamber set-up which also provided the vacuum pump

mentioned previously. The gauges themselves were found to be quite accurate in the range of pressures 10-1000μm Hg.

3.6 Data Acquisition System.

A schematic of the data acquisition portion of the experimental apparatus is illustrated in Figure 3. Pressure, temperature, and heat flux voltage outputs were amplified as necessary, filtered of undesired high frequency noise, and processed by a Masscomp 5600 computer. This machine was capable of sampling beyond 0.3 MHz and had a 4 Megabyte RAM. As a result, we were by no means limited by our digital hardware. The Masscomp was connected in such a way as to be able to read eight inputs essentially simultaneously. Though the sampling rate for each experiment was 1 Hz, the burst frequency, or the inverse of the sampling time between each channel, was set at 1000 Hz Burst frequencies above this value yielded identical results during a number of preliminary tests. This was most likely due to the response time of the thermocouples being of the order of 100 Hz at best. The assumption of "instantaneous" measurements was therefore substantiated.

3.7 Pressure Measurements.

Transient and steady pressures were measured using pressure transducers whose voltage output was a linear function of the pressure in question. The transducers used in this experiment were all tested for linearity and were re-calibrated with a mercury manometer to improve on the accuracy determined by the manufacturer. The transducers were equipped with temperature correction circuitry up to 85°C so compressed air cooling lines were required for each transducer during operation of the system.

3.8 Temperature Measurement.

Type-T 36 gauge thermocouple wire was used for the majority of the temperature

measurements, except for in the cases of the heat flux transducer calibration experiments where 0.003″ diameter wire was used for improved time response, and for ambient temperature measurements where a larger thermocouple could be accommodated. All thermocouple leads and connectors were checked for homogeneity.

An ice bath was incorporated into the thermocouple circuit to avoid the inaccuracies associated with electronic ice baths. The reference thermocouples were sealed in separate oil filled test tubes and inserted into the bath to minimize any irregularities contributed by convection currents in the ice bath itself. The bath reference temperature was checked before each experiment, which did not progress until this temperature reached 0.0°C precisely.

To improve the accuracy of the temperature measurements, the output voltage verses temperature data published by the manufacturer for the Type-T thermocouple was refitted with a polynomial curve fit routine over the specific temperature range of interest. This procedure reduced the curve fitting error from ±0.5°C to ±0.1°C for the range 15-160°C. The hardware error for the temperature measurements was thus ±0.1°C. The details of this fit are provided in the Appendix (A1). Correction calculations needed to be performed separately to incorporate conduction, convection, and radiation losses. Most temperature measurements needed not to be corrected at all.

3.9 Heat Flux Measurements and Calibration.

A number of heat flow sensors called Microfoils (heat flux transducers) were purchased from The RdF Corporation for the purpose of monitoring the energy transferred to the spheres during the experiment. These instruments, much like strain gauges in appearance, utilized a pair of thermopiles separated by 0.001″ in order to output a voltage proportional

to the heat flow through it. The dimensions of each transducer were $0.25''$ by $0.70''$, the area of the actual heat flow measurement being roughly $0.25''$ by $0.35''$.

The voltage output by these devices was of the order of μVolts which necessitated added amplification. The material used as the substrate in which the thermopiles were imbedded was Capton, a nonconductor with thermal conductivity of $0.118 W/m^{2^\circ}$C. Since the thermal impedances of the heat flux transducer substrate and the sphere wall (304L SS; $k = 14.9 W/m^{2^\circ}$C) and PCM combination differed greatly, separate calibration experiments needed to be conducted in order to determine a multiplication constant for the linear output of the gauge. The multiplication constants include the contact resistance contributed by the glue used to attach the sensor to the spheres. In most cases this glue was Plastic Steel, an epoxy.

The thermal conductivity of the substrate material for each heat flow sensor was a slight function of temperature. To correct for this added complication, a fifth order polynomial curve fit was implemented which used, for a data base, the temperature variation of the conductivity of the substrate provided by the manufacturer of the sensor. Each sensor was received with individual specifics concerning μVolt output per heat input, response time, dimensions, and dependence on temperature data.

Two means of calibrating the transducers were devised. The first employed a comparison between the heat flow sensor and a thermocouple pair measurement technique for a thick stainless steel plate exposed to a condensing steam jet. Experiments of this nature were also performed on the spheres, but were less conclusive because of the small shell thickness which made temperature measurements across the wall difficult.

In these tests, an intrinsic thermocouple pair, one on either side of the $0.25''$ stainless

steel plate, was attached next to a heat flux transducer. A steam jet was then directed at the plate and the output of the transducer and thermocouples were recorded in time by the Data Acquisition system described earlier. A schematic of this approach is provided as Figure 4. The temperature values output by the thermocouple pair were then fed into a numerical routine which employed a Crank-Nicolson scheme to solve inversely for the instantaneous heat flux. These values were then compared to the outputs of the flux transducer and a multiplication constant (MF) was calculated, based on the integrated heat flux values (or energy transferred). Typical plots of these experiments are provided as Figure 5 and 6, the latter presenting the results of the former multiplied by an experimentally determined constant—in this case, MF=1.1. The software used for this calibration procedure was verified to produce extremely accurate results when compared to analytical solutions. A brief equation formulation and numerical to analytical comparison is included in the Appendix (A2).

The method described above was effective for accurately obtaining a multiplication factor, but since the heat transferred into the PCM was also a function of the wall thickness and a moderate function of the heat transfer rate†, the constant arrived at through these means needed to be scaled accordingly. The method of scaling was simplistic, incorporating a ratio of the resistance of the wall to the resistance of the wall and heat flux transducer itself. Using the experimental values for the thick plate experiment, only a range of multiplication factors could be determined, between 1.1–8.5. These calculations are included in the calculations section of the Appendix (A3).

Another strategy needed to be adopted to find the multiplication factor with a bit

† The heat flux transducer's thermal capacitance adds a complication when the device is used for transient measurements. Preliminary experiments have shown a slight influence due to the transients of the sensors themselves. Quantification of these effects will be soon in coming.

more accuracy. A second means for this task was to ratio the summed heat fluxes over time, output by the heat flux transducer, to the analytically determined total energy considering the thermodynamic initial and end states. It was found that the heat flux into the larger spheres, whether isolated or in an arrayed configuration, was predominantly uniform, less so for the 2.0″ spheres as will be explained. This was most likely due to the solid PCM remaining suspended in the center of the sphere around the probe used to measure its center temperature. Gravitational effects on the heat flux distribution were therefore reduced making comparison to a previous study less straight forward (Moore and Bayazitoglu, 1982). Even when examining the most extreme cases, for a given arrangement, the heat flux histories appeared equivalent only differing by a certain amplification factor. This meant that though the heat flux varied slightly over the surface of a given sphere, the overall heat flux history of the sphere could be traced-out by multiplying the experimental energy transferred by the inverse ratio of the total experimental energy stored to the total possible thermodynamic energy storable given an initial and final state—both in temperature and phase. The multiplication factors used in this study were determined using this technique and are listed with each plot of results. One can note these values are bounded by the range found experimentally by the previous method, namely 1.1–8.5. Errors of these value were typically ±4%. A sample calculation for the total energy transferred is also presented in the Appendix (A3).

3.10 Spheres-Filling and Instrumentation.

The PCM encapsulating spherical shells were manufactured by Seymour-Sheridan, a metal floats producing company out of Connecticut. Each shell consisted of 304L Stainless Steel. The PCM selected for this series of tests was P-116, a paraffin, provided by the

Sun Oil Company. The relevant thermophysical properties for each material are listed in the Appendix (A4).

The spherical shells themselves varied in thickness depending on the diameter of the particular sphere. These thickness's were 0.020, 0.024, and 0.040 inches for the 2.0, 2.5, and 3.0 inch diameter spheres respectively. Each sphere was fitted with a $1/4X20$ thread to mate with 0.25 inch long 304 Stainless machined screws which provided for time efficient processing of the spheres and for material continuity.

All of the spheres were filled in basically the same manner. The non-instrumented spheres were injected with liquid PCM at about 120°C until the fluid brimmed the sphere. The stainless screws, wrapped with teflon tape, were then inserted into nylon washers and tightened down into the fittings of each sphere. The warm nylon washer when squeezed between the screw head and the sphere provided a leak proof seal. Since the spheres were completely filled with liquid PCM, a relatively high vacuum (at least $100\mu m$ Hg) was drawn upon solidification of the wax. If the seal was not perfect, air would be sucked into the shell which would then result in a pressurized sphere during the next melt cycle. Random spheres were thus leak tested in hot oil and observed for escaping air. No leakage was noticed for oil bath temperatures 50°C beyond the liquid fill temperature. Spheres weighed before and after particular experiments showed no PCM loss which has been a problem in a past study by Graves (1985).

The instrumented spheres were treated similarly as the non-instrumented spheres only a special screw was fabricated so as to permit the temperature at the center of the sphere to be measured. A schematic of this operation is included as Figure 7. Holes drilled in the stainless steel screws at 90 degrees from each other allowed for the fine

thermocouple wire to pass through and into the center of the sphere. Hollow ceramic dowels glued to the screw tips guided the thermocouples directly to the center of the spheres, where they remained for the duration of each test regardless of the churning effects of the melting process. Sealing against leakage for these spheres was more difficult than for the non-instrumented spheres, yet eventually a "perfect" seal could be achieved for each sphere.

The dowels used to measure the center temperatures acted to even out the heat flux in the tangential gradient. Figure 8 illustrates this procedure. Normally, the PCM solid would settle to the bottom of the sphere where it would create a non-uniformity by increasing the heat flux over the lower portion of the sphere via direct contact between the hot shell wall and PCM solid. The center temperature probe prevented this from occurring by maintaining the solid at the center of the sphere. Gravitational effects were therefore limited to natural convection in the liquid PCM. This smoothing out of the heat flux over the sphere enabled single heat flux transducer measurements to be reasonably representative of the average heat flux to the sphere in time.

4. EXPERIMENTAL PROCEDURE

As stated in the original proposal, the purpose of this primarily experimental research effort was to monitor quantitatively in time the significant heat transfer parameters during the dual latent heat energy exchange between a condensing vapor and a packed bed of melting PCM encapsulated by a metallic spherical shell. The apparatus described above was designed to be well integrated enough to produce a controlled, repeatable environment with which this could be accomplished.

The main objective was to be able to produce a clean, noncondensible gas-free environment which could as rapidly as possible provide a steady *ambience* for an isolated PCM-filled sphere, or array of spheres, in terms of pressure, temperature, and velocity (or mass flow rate), so that the energy transfer characteristics of the process could be observed. What is meant here by *ambience* is condensible vapor, in this study—pure steam.

The thermal loop presented incorporates a "dummy" loop, or dummy chamber approach, in order to establish certain prescribed thermodynamic conditions. Once the desired conditions were achieved in the dummy loop they could be easily transferred to the "real" loop which housed the spheres of interest. The dummy chamber needed only to be monitored for temperature and pressure.

4.1 Preparatory Procedure.

A number of steps to ready the system for the actual data taking were regularly preformed before each run. The first step was to rid the entire system of noncondensible gases. This was initially accomplished by evacuation, but once the components got "wet" this approach became incredibly time consuming. As a short cut, degased water was pumped through the bypass line with the system in the open loop configuration. Not only did the water force out a large portion of the trapped air, but the degassed water, with its affinity for regasifaction, also assisted in purging the system of noncondensibles. Evacuation was reserved for the chambers and began a minimum of two hours before each test.

While the fluid was circulating through the bypass loop, the preheater and boiler were turned on until saturated steam at slightly above atmospheric conditions was generated.

(This also helped to guarantee a minimum of noncondensibles in the fluid transport lines of the thermal loop.) The valve to the bypass loop was then closed and the inlet valve to the dummy chamber was opened. The chamber was allowed time to pressurize until it too reached slightly elevated atmospheric saturation conditions. The exit valve of the dummy chamber was then opened which enabled the fluid to circulate again. Dummy chamber temperatures and pressure were monitored visually by use of digital thermometers and voltmeters.

While steady conditions were being reached in the dummy chamber, evacuation continued on the real chamber. The real chamber pressure was typically around 380μm Hg (or 5% of 1% of an atmosphere) for the initial tests, but it was found that after numerous tests the condensate remaining in the chamber became significant and evacuation to these low pressure levels was difficult. It was assumed eventually that most of the pressure in the chamber was created by the presence of water vapor and not air and that extensive evacuation was not required. As a result, the real chamber was evacuated for a set period, roughly two hours. Vacuum pressures attained in this manner were generally 1000μm Hg.

4.2 Inlet Conditions Technique(s).

Once the circulating fluid reached and maintained steady conditions in the dummy chamber loop, a number of techniques could be used to "transfer" these conditions to the real chamber and the spheres. The technique selected depended on the initial condition desired for the vapor inlet condition to the sphere ambience. One strategy required preheating the real chamber in order to rapidly come to steady chamber pressure and temperature. Unfortunately, this approach led to heat transfer to the sphere via radiation before the experiment was initiated. This was most undesired. The second attempt to

achieve rapid steady flow conditions in the real chamber was to adopt a burst method. This method involved closing the exit valve to the dummy chamber once it had reached the steady flow condition. Steam pressure in the dummy chamber was allowed to build to a prescribed value whereupon the inlet valve to the real chamber was opened. At this time, not only did the real chamber receive vapor generated by the boilers, but also the excess vapor stored in the previously pressurized dummy chamber. The dummy chamber entrance valve was closed at another predetermined pressure, while the exit valve of the real chamber remained closed so that build-up to the steady conditions could proceed. Once this condition was reached the real chamber exit valve was opened and the flow was established, now in the real chamber loop. The inlet transport line from the boiler to the real chamber was preheated to a temperature slightly higher than the steady flow condition temperature. This process had a negligible effect on increasing the temperature of the real chamber before the test and helped to guarantee that no condensate would enter the chamber. The data sampling began normally 30 seconds before the inlet to the real chamber was opened. The heat flux measurements, relevant temperatures, and pressures were sampled every second for 30 minutes. Evacuation of the real chamber was terminated less than 5 seconds before the valve to the real chamber was opened.

This "burst" technique preheated the real chamber, in a sense, by dumping a large quantity of condensing vapor into it. The technique was quite easily repeatable as will be seen, but had the drawback of not being able to produce steady flow conditions as fast as the first method mentioned which employed four heating tapes to preheat the chamber. The burst technique was used in each of the experiments performed in this study. A third technique involved radiation shielding to reduce the heat transferred to

the sphere(s) during the preheat process using the heating tape. This technique was found to be superior to the two discussed so far, but was not developed until recently and could not be used without repeating all experiments. The details of the burst method are listed in the Appendix (A5).

4.3 Experiments.

Each experiment involved a certain sphere (or sphere configuration), instrumented with a heat flux transducer, a sphere surface thermocouple, and a center sphere thermocouple, which was filled with solid PCM and placed in the real chamber. Opening the inlet valve to this chamber began the energy transfer process between the condensing steam and the melting PCM. For these tests, 50% of the energy transferred to the spheres was latent due to the subcooled solid PCM at the initiation of the each experiment and the excess heating of liquid PCM above its melt temperature. The PCM's melting point was 45°C at atmospheric pressure. The initial and final temperatures of the spheres for each run were generally 20 and 105°C respectively. The percent of total latent heat storage can be seen by observation of the example total energy stored calculation provided in the Appendix (A3).

In order to be able to develop a better understanding of the effects of neighboring spheres in a packed bed experiencing dual latent heat transfer, a sort of comparison approach was employed. This approach entailed gathering data for single isolated spheres of varying diameters and comparing it to the data for the same spheres in an array of other spheres of the same diameter, holding the inlet conditions constant.

4.4 Sphere Arrangements for Comparisons.

As mentioned, three different spheres sizes were examined. For each sphere size a

number of single sphere and arrayed sphere tests were performed. For the single sphere tests (ex. take the 3.0″ diameter sphere) the appropriate sphere holder (in this case the 9X9 sq. inch aluminum insert) was inserted into the real chamber and an instrumented sphere was suspended by a rigid insulated wire in the center of the chamber. The chamber was then sealed, evacuated, and the test was performed.

To get an idea of the bed effects, the same instrumented sphere was retested, only for these tests the sphere of interest was located in the center of a 3X3X3 array of spheres of the same diameter (refer to Figure 2). The stacking arrangement was that of the (1,0,0). These tests were conducted identically to the single sphere tests, the ambient pressure and temperatures being measured at the inlet to the aluminum insert, and the heat flux, wall temperature and center temperature data being recorded in time.

5. EXPERIMENTAL DATA AND INTEGRATED FLUX RESULTS

Typical data resulting from these tests are presented graphically in Figures 9-14. These plots reveal the heat flux and temperature histories for the single and arrayed sphere configurations for each sphere diameter. Figure 15 is provided for comparison to Figure 14, and Figure 16 for comparison to Figure 9, to show the repeatability of the data and the measurable irregularities in the heat flux associated with the unsteady heat transfer process. Figures 17-22 plot the integrated heat fluxes output by the heat flux transducers against time for sphere diameter and sphere configuration. Figures 23 and 24 are provided as checks for Figures 22 and 17, respectively. Also included on these plots is the time dependency of a "bulk" heat transfer coefficient which was calculated using the

temperature difference between the ambient vapor and the outside surface of the sphere.†

The same steady flow condition was achieved in the dummy loop for each experiment, and the same burst technique was used to inlet the condensing vapor to the real chamber, and hence, the sphere ambience. Example plots of the ambient temperature and pressure for two distinctly different runs of the 3.0″ sphere are provided as Figures 25 and 26. The first figure shows a relatively rapid achievement of steady conditions while the second illustrates a more smoothed ambient response to the same inlet condition as the former. The difference in ambient temperature and pressure is attributable to the larger heat sink of the arrayed sphere configuration of Figure 26. This sink allows more vapor to condense therefore resulting in a longer time to reach the steady flow conditions of the dummy chamber. For the arrayed sphere configuration, steady inlet conditions are maintained for only the latter 30% of the energy exchange process. These figures provide the most distinction between ambient conditions for the single and arrayed sphere tests. The others are less dramatic because the heat sinks are proportional to the volumes of the spheres.

Figure 27 is provided to illustrate the repeatability of the ambient conditions. As compared with Figure 24, these ambient temperature and pressure histories are practically identical, the only differences can be associated with slightly different begin times and initial temperatures.

6. RESULTS AND DISCUSSION

† These heat transfer coefficient values are most useful in revealing general trends. The actual values are disputable based on the fact that the sphere outside surface temperature was measured by a thermocouple imbedded in the heat flux transducer substrate, and that, at least for the arrayed tests, the ambient temperature was not measured locally, but as stated, at the entrance of the aluminum inserts. These two experimentally introduced inaccuracies may have had some cancelling effect since the former would tend to decrease the coefficient while the latter would increase it.

The energy storage data for each test is condensed in Figure 28. In this plot the instantaneous energy storage is divided by the total thermal capacities determined by the respective initial and final thermodynamic states and are plotted in time. For each sphere size and configuration, the 95% energy storage times are provided which give some indication of the differences in heat storage rates at a glance.

6.1 Data Interpretation.

The same procedure was used for each test–only the spheres and aluminum inserts were changed during the experiments. Unfortunately, the varying heat sink potential for each sphere size caused a resulting effect of the ambient condition as seen by comparing Figures 25 and 26. Therefore, each sphere was exposed to a different ambient condition, the greatest differences being between the single and arrayed sphere configurations. Comparisons can be made quite readily, however, if limited to either just isolated spheres, or just arrayed sphere configurations. Figure 29 and 30 are provided for this purpose. From these figures one can see more clearly that, as expected, the energy storage histories take the form of the error function in descending order according to the diameter of the sphere. One encouraging feature in Figure 27 is the large linear portions of the energy storage curves which have been calculated in a previous numerical study by Moore and Bayazitoglu (1982). These smoother profiles indicated that one of the effects of the bed, or neighboring spheres, is to provide a more uniform heat transfer process. This is also visible from Figures 10, 12, and 14, which show a more constant heat transfer coefficient. Figure 14 is the best example of this when compared to Figure 13. This finding may dramatically simplify dual latent packed bed design calculations, which could employ constants rather than time dependent variables. This extended linear portion of the

energy storage curves is not as prevalent in the single sphere tests, especially for the 2.0" diameter sphere.

As previously mentioned, the heat flux to the sphere was kept relatively uniform over its surface area by the fact that the PCM solid remained suspended in the center of sphere. This aspect of the heat exchange process may provide a good indication for the zero gravity condition of space by limiting the gravitational effects to natural convection in the molten PCM. Unfortunately, this "suspension" could not be maintained for the entire duration of each test. As seen in a glass sphere test at a certain time late in the experiment, the PCM solid would slide off the ceramic dowel resulting in the center temperature histories of Figures 9-16 and the corresponding jumps in the heat flux from transducers attached to the lower hemisphere of the sphere. Though repeatable, this phenomena was undesirable, and could be eliminated by orienting the spheres with the probes extending horizontally, but time was not available to repeat the experimental set. One interesting feature demonstrated from this occurrence was that the heat flux between the solid PCM and the shell wall could be up to five times that of the heat flux between the PCM liquid and the shell wall (see Fig.s 9 and 16 heat flux curves). Space systems may simply differ from earth systems by a scaling factor without too much loss of accuracy.

Most of the wavy nature of the 2.0" sphere data (Fig. 28) can be explained by the above discussion concerning the release of the PCM solid to the bottom of the spherical encapsulate. This effect is reduced for the larger spheres because, since the size of the solid PCM at the time of release is basically constant for each sphere, the effect is diminished as the sphere volume is increased as seen in figures 11–15.

The heat transfer coefficients calculated and plotted with the energy storage histo-

ries are bulk quantities in the sense that they include all latent effects. The experimental method to calculate the heat transfer coefficient was simplistic. Once the heat flux was know for a particular time the ambient and wall temperatures were differenced and used directly to calculated h. The values for h determined in this manner were relatively constant during the major portion of the exchange process and ranged around $100 W/m^{2\circ}C$ for each sphere. The sphere wall temperature was measured using the thermocouple embedded in the heat flux transducer substrate. This experimental procedure led to an awkward temperature correction which was eventually abandoned. Left untreated, the outside wall temperature measured by the transducer thermocouple over estimated the actual wall temperature. In some cases this difference was as high as 2.0°C. The effect of this reading was to calculate h's based on a smaller temperature difference, meaning; the heat transfer coefficients calculated using experimental results were likely higher than they should have been. (The other effect was to produce the intense oscillation in the steady end states of the experiments; see plots.)

6.2 Condensed Data and Scale Analysis.

As seen from Figure 28, the percent energy storage parameter (ϵ) is well behaved and should be since its scale can be clearly defined if the initial and final thermodynamic states can be known, namely, $E_{scale} \sim m_{pcm}(h_{fg} + \bar{C}p\Delta T)$.† This guarantees that ϵ, a dimensionless quantity, will abound between 0 and 1. This range is essentially known exactly because E_{scale} may be accurately determined.

Scaling the real time component, t, is not that straight forward, however. The heat

† From this point, barred quantities will represent any "phase change compensated" quantity: ex. $\bar{C}p = (Cp_l\Delta T_l + Cp_s\Delta T_s)/\Delta T$, where $\Delta T_l = T_f - Tm$, $\Delta T_s = Tm - T_i$, $\Delta T = T_f - T_i$, and l, s, i, f, m denote liquid, solid, initial, final, and melt respectively.

transfer process for our case is complicated by the multiple resistances and their respective time scales. For brevity, an elaborated development will be avoided here, but it should be noted that the method is physically motivated, derived from the energy equation and includes transient conduction, convection, and latent heat terms.

For our case the time scales for the heat transfer across the condensate film (mostly dropwise anyway) and the encapsulating stainless shell ($\alpha_{shell} \gg \alpha_{PCM}$; where α = thermal diffusivity) are negligible with respect to the major scales of the heat transfer across the liquid PCM to the solid and the heat transfer in the solid itself. The former of these second two primary resistances to heat exchange "undercuts" the latter in that the solid PCM never really is allowed to store energy because the melt front (in our system) moves faster than the heat wave through the solid PCM. The effect can be described as distributing the spheres energy across the molten PCM, or in lowering the sphere Biot number. With this considered, a resulting time scale for the phase change process may be determined. For our system, properties, and dimensions this scale is

$$t_{scale} \sim \frac{r^2}{\alpha_{eff}}(1 + \frac{h_{fg}}{\bar{C}_p \Delta T}) \sim Fo_{eff}(1 + Ja^{-1}) \qquad (1)$$

where r=sphere radius, $\alpha_{eff} = k_{eff}/\bar{\rho}\bar{C}_p$, and h_{fg} = latent heat, k, ρ, and C_p being the thermal conductivity, density and specific heat of the PCM, respectively. This is the most simplified form. Normally more, addend-type terms would appear on the end of equation 1, but for all practical purposes, these values add little and may be neglected.

The real time scaled (divided) by this time (t_{scale}), which is really just a prediction of how long it will take to transfer the energy, should be positive and $O(1)$. A good check for this t_{scale} is to disregard the phase change and assume purely sensible storage. The

latent term in the scaled result drops out and k_{eff} becomes k of the PCM, or

$$t_{scale} \sim \frac{r^2}{\alpha} \tag{2}$$

which is well known as being the appropriate scale for solid spherical heat transfer (Bejan, 1982). The effective conductivity, k_{eff}, appearing in equation 2, accounts for the enhanced conductivity of the liquid due to internal circulation. It is really not k_{eff} that should be discussed, but rather k_{eff}/r, which is essentially the internal heat transfer coefficient scale. One can see that in a zero gravity environment $k_{eff} = k$, a constant.†

The percent energy for the arrayed spheres, ϵ, is replotted against the new dimensionless time ($\tau = t/t_{scale}$) in Figure 31. The collapse of the data into a single curve for the entire range of time is astounding. The magnitude of τ being of $O(1)$, in the figure roughly 0.6, also suggests an accurate selection of parameters for t_{scale}. The collapse is not complete for the $2.0''$ sphere which was shown earlier to be significantly effected by the sliding off of the PCM solid during the charging process. This curve, though paralleling the others to some degree, should not be considered too seriously since the "sliding" effect had not nearly the significance for the larger spheres as for the $2.0''$. Regardless, the meaning of Figure 31 could be profound. With this information one would only need to know the end states, material properties and dimensions to know the percent energy storage at any time. This result, of course, is most applicable to a single sphere than a bed of spheres since the Reynolds number for these experiments was extremely low—$O(10)$. The condensed curve in Figure 31 is thus most applicable for dual latent thermal energy storage with stagnant condensing vapor flow condition.

† The value of k_{eff} in this development was determined experimentally to be $2.8 W/m^{\circ}C$, nearly 18 times the value of the conductivity of the liquid phase without internal convection.

7. RECOMMENDATIONS FOR FUTURE STUDY

It is apparent from these results that this work is still in its adolescence. Some enlightening findings were discussed above which could not have been possible without the spinoffs of the trial and error nature of experimentation.

A marked increase in the comparability of the data could be made by the following adaptations of the experimental procedure:

1. The burst method discussed in the procedures section could easily be replaced by the new technique which uses a preheat method of the real chamber using heating tapes with radiation shields which limit the pre-experiment heat transfer to the real chamber itself and not to preheating the spheres. The adoption of this technique would provide far more rapid steady inlet conditions than were achieved in this study. In addition, for the single sphere tests, identical ambient conditions could be created as for the arrayed sphere tests by suspending the single sphere up stream from the multiple sphere arrangement. The arrayed spheres would act in this sense as a heat sink in order to duplicate the ambient conditions of the arrayed sphere tests. This would provide more of a basis for comparison between the single and arrayed sphere configurations.
2. The next improvement which could be easily made is to rotate the sphere in such a way so that the center sphere temperature probe is horizontal. This would maintain the PCM solid at the center of the sphere until it was nearly, or completely, melted. Another means to accomplish the same ends would be to fabricate a barb network about the end of the probe so that regardless of the orientation of the sphere, the PCM would remain suspended. This barb could be made of a nonconductor with

low thermal properties which would minimize its influence in the heat transfer process.

3. Lastly, the outside surface of the encapsulating shells for each instrumented sphere could be fitted with an intrinsic thermocouple. This accurate wall temperature would benefit the precision of the heat transfer coefficient calculations made by experimental values. This in turn, would promote better comparisons to theoretically determined heat transfer coefficients for the case of transient condensation on a spherical shell containing a PCM. More local ambient temperatures could be used for these calculations which begins to extend the reach of this project to the actual packed bed level.

8. CONCLUDING REMARK

The procedure and results presented here reveal a steady repeatable environment in which control is provided enough to make reasonable comparisons with previous and future analytical and numerical predictions. What is suggest by these recommendations is that improvements in the experimental procedure will make such comparison more clear without the hand waving arguments associated with order of magnitudes, etc.

Each of the proposed steps listed above will be included in a thesis which should be completed by mid-May. Also to be included will be 1) a more quantitative dependence of the heat flux as a function of the tangential angle while the PCM solid is suspended in the center of the sphere, 2) a complete study on the transients of the heat flux transducers themselves, and 3) a small scale packed bed experiment including the axial dependence on the energy storage rates of a bed $3X3X8$ spheres in dimension. Wall effects will be

observed by repeating certain experiments performed on a $3X3X3$ network to a $5X5X5$. We expect our findings will be beneficial to you and that extension of our experimental approach, now in its finalization stages, to higher temperatures and higher temperature materials will solicit your interests and support.

REFERENCES

Amundson, N.R., 1956, Solid-fluid interactions in fixed and moving beds. Fixed beds with small particles, Ind. Eng. Chem. 48, p. 26.

Beasely, D.E. and Clark, J.A., 1984, Transient response of a packed bed for thermal energy storage, Int. J. Heat and Mass Transfer, 27, p. 1659.

Bejan, A., 1984, Convection Heat Transfer, Wiley-Interscience publication, p. 19.

Chung, J.N., 1986, A numerical simulation of the liquid-metal dual-latent heat packed bed thermal energy storage system, Final Report to AFOSR under 1986 USAF-UES Summer Faculty Research Program.

Graves, A.G., 1985, Transient Thermal Performance of an Experimental Packed Bed Thermal Energy Storage System, a Masters Thesis, U. of Tenn.

Gross, R.J., Hickox, C.E., and Hackett, C.E., 1980, Numerical simulation of dual-media thermal storage systems, J. Solar Energy Eng., 102, p. 287.

Jones, D.E. and Hill, J.E., 1979, Testing of pebble-bed and phase-change thermal energy storage devices according to ASHRAE Standard 94-77, National Bureau of Standards, Report NBSIR 79-1737, WA D.C.

Kordahi, M.E., Yimer, B., and Crisp, J.N., 1982, Transient thermal analysis of phase-change thermal energy storage system with density change, ASME paper, 82-HT-10.

Martin, H., 1978, Low Peclet number particle-to-fluid heat and mass transfer in packed beds, Chem. Eng. Sci., 33, p. 913.

Robertson, G.F. and Cavendish, J.C., 1981, The effects of flow maldistribution on the thermal response of packed bed converters and thermal energy storage units, ASME paper, 81-HT-59.

Peavy, B.A. and Dressler, W.E., 1977, Transpiration heat transfer in thermal energy storage devices, National Bureau of Standards, Report NBSIR 77-1237, WA D.C.

Shitzer, A. and Levy, M., 1983, Transient behavior of a rock bed thermal storage system subjected to variable inlet air temperatures: analysis and experimentation, J. Solar Energy Eng., 105, p. 200.

Schumann, T.E.W., 1928, Heat Transfer: A liquid flowing through a porous prism, J. Franklin Inst., 208, p. 405.

Torab, H. and Beasely, D.E., 1985, Dynamic response of a packed bed of encapsulated PCM, Proceedings of the 20th Intersociety Energy Conversion Eng. Conference, Miami Beach, FL, p. 313.

Yimer, B., Crisp, J.N., and Mahefkey, E.T., 1980, Transient thermal analysis of phase-change thermal energy storage, ASME paper, 80-HT-2.

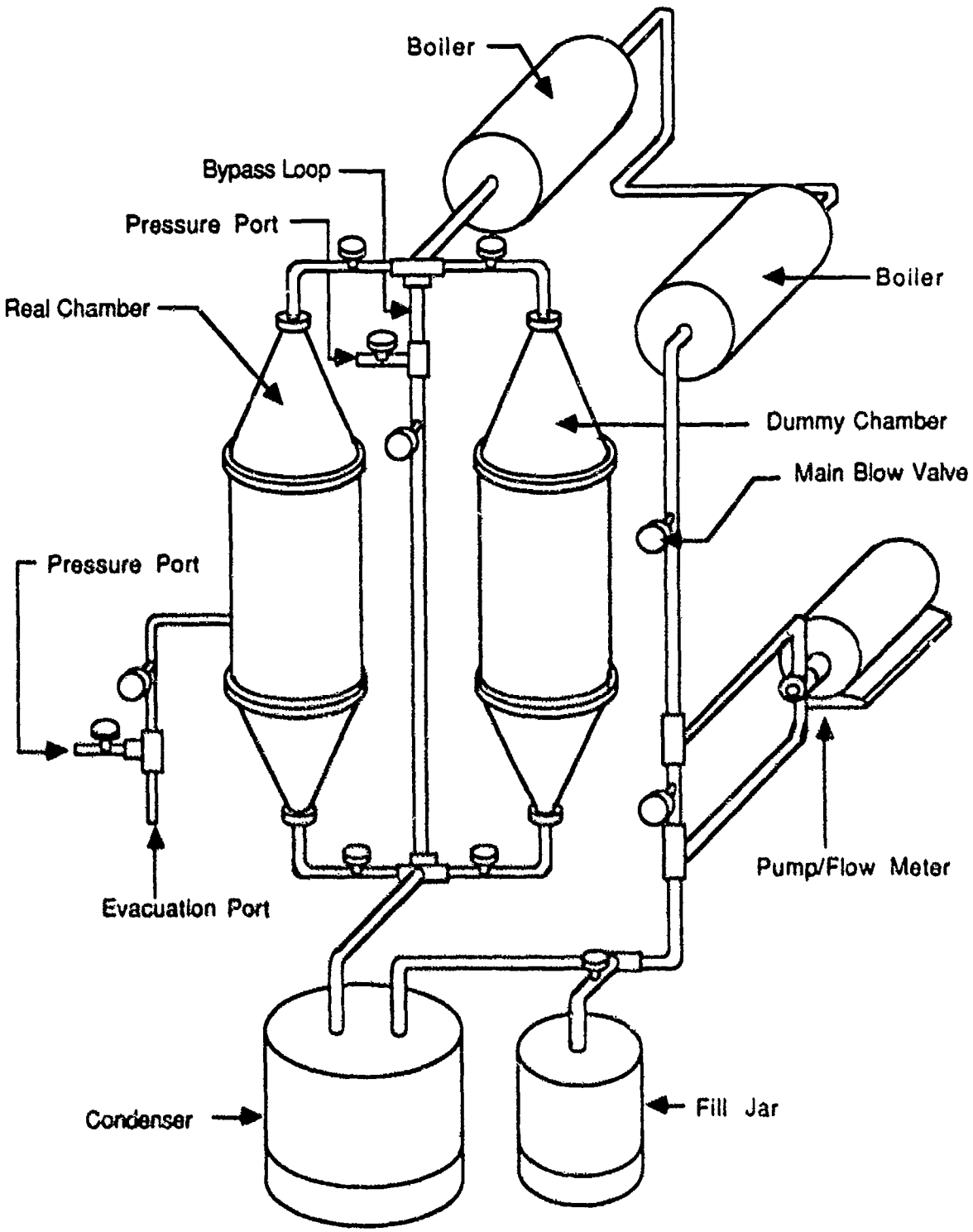


Fig. 1. System Schematic

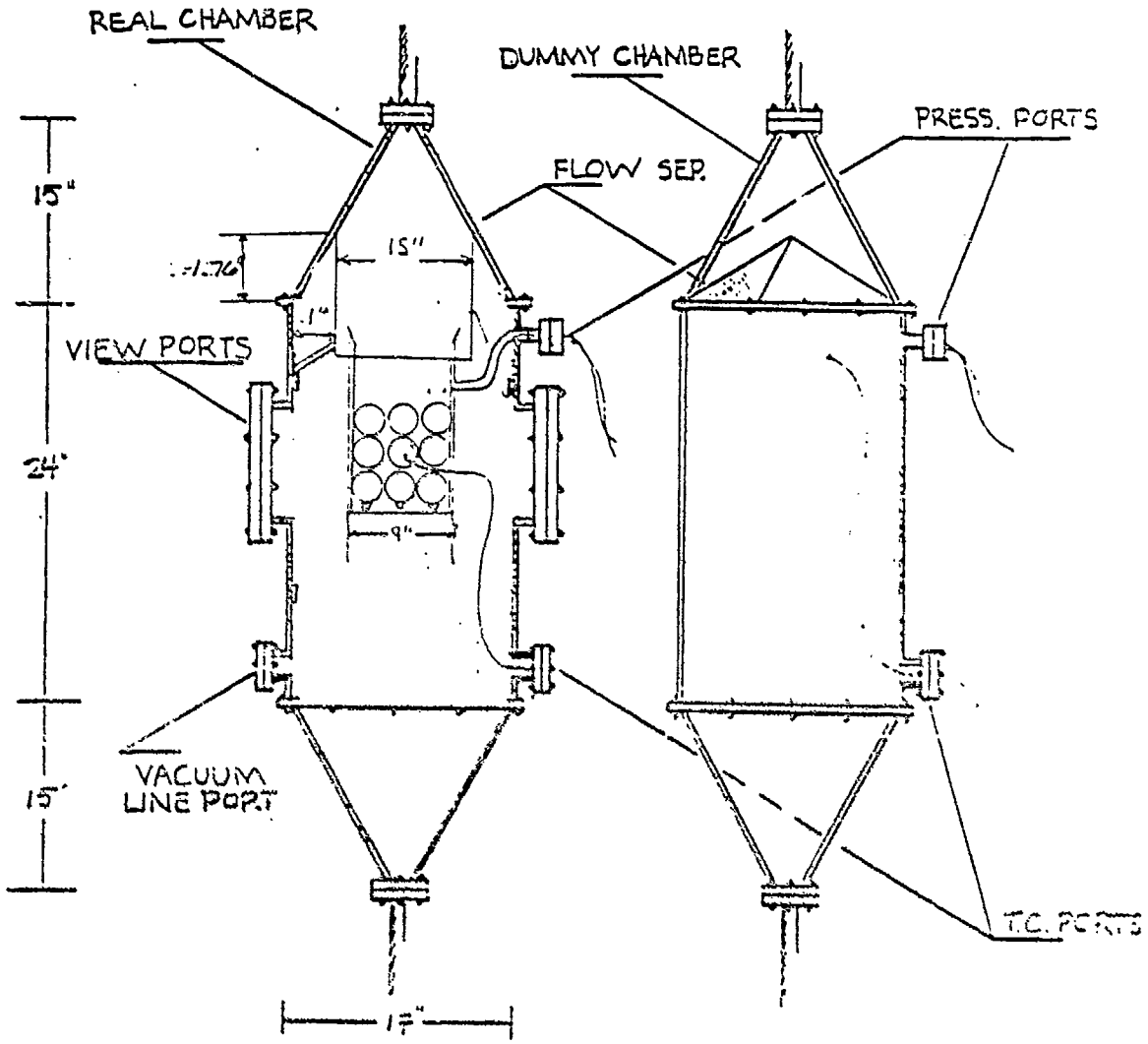


FIG. 2 GENERAL SCHEMATIC OF TEST CHAMBERS [DIM. IN INCHES]

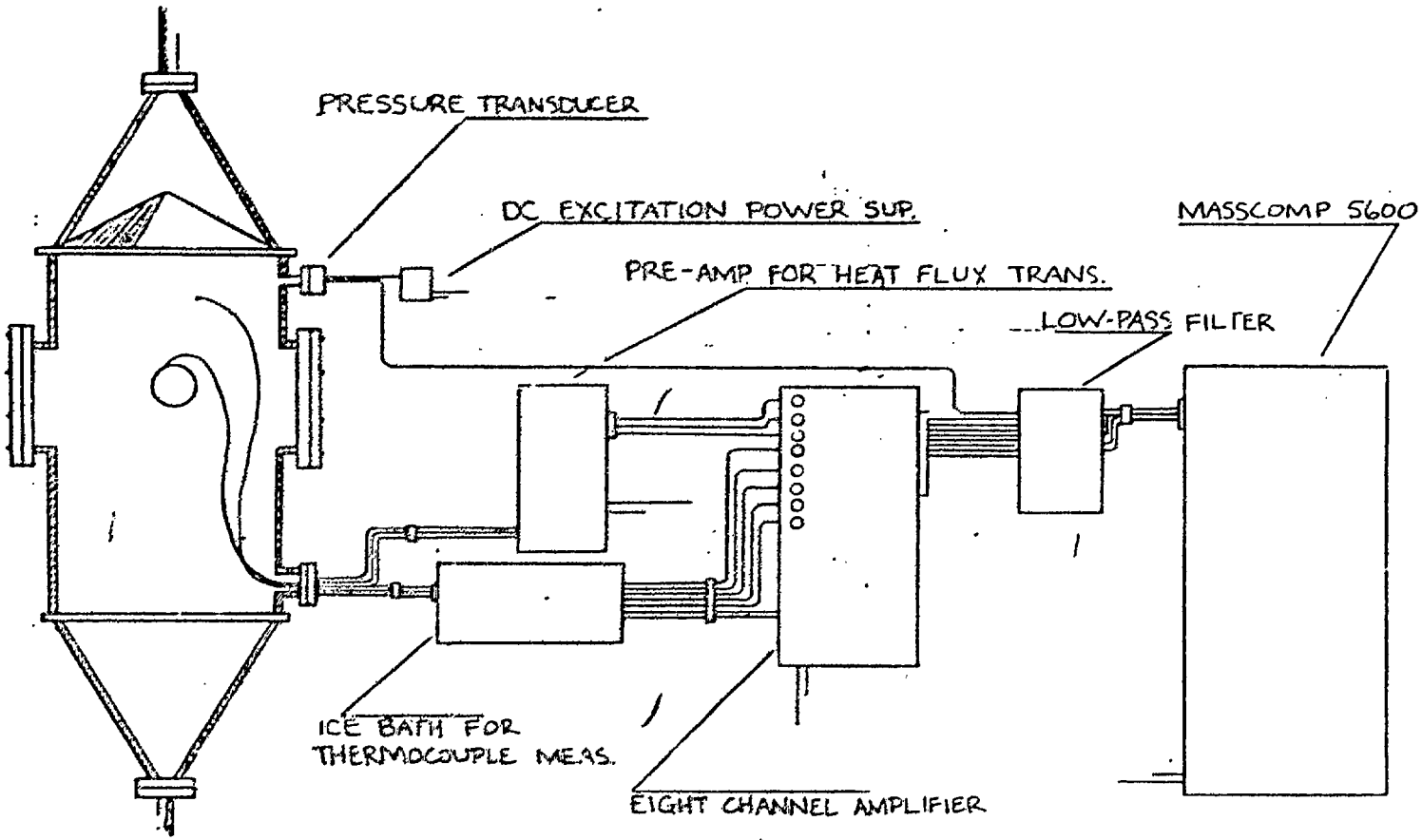


FIGURE 3 SCHEMATIC OF DATA ACQUISITION SYSTEM W/ REAL CHAMBER.

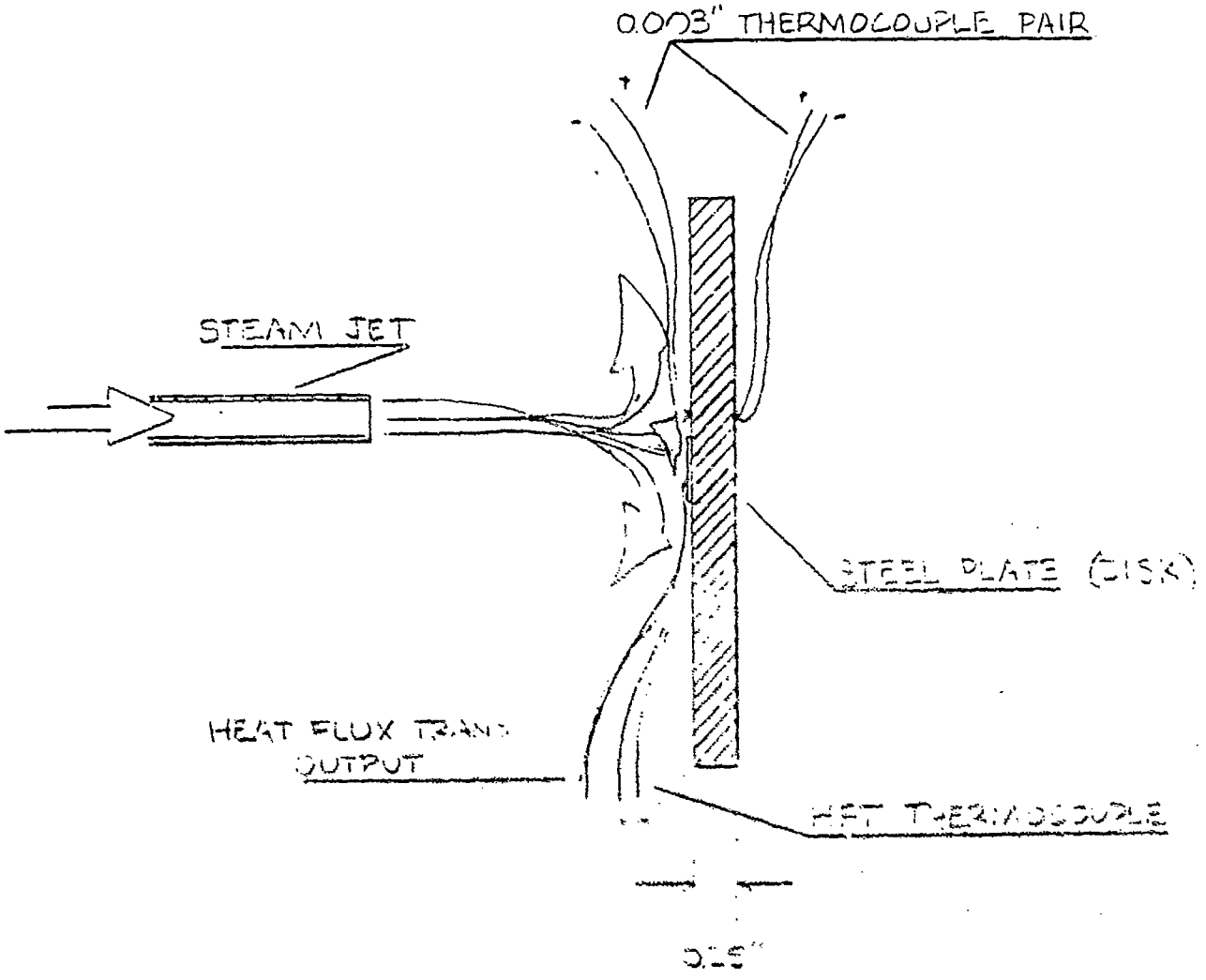


FIGURE 4 SCHEMATIC OF HEAT FLUX TRANSDUCER CALIBRATION EXPERIMENT.

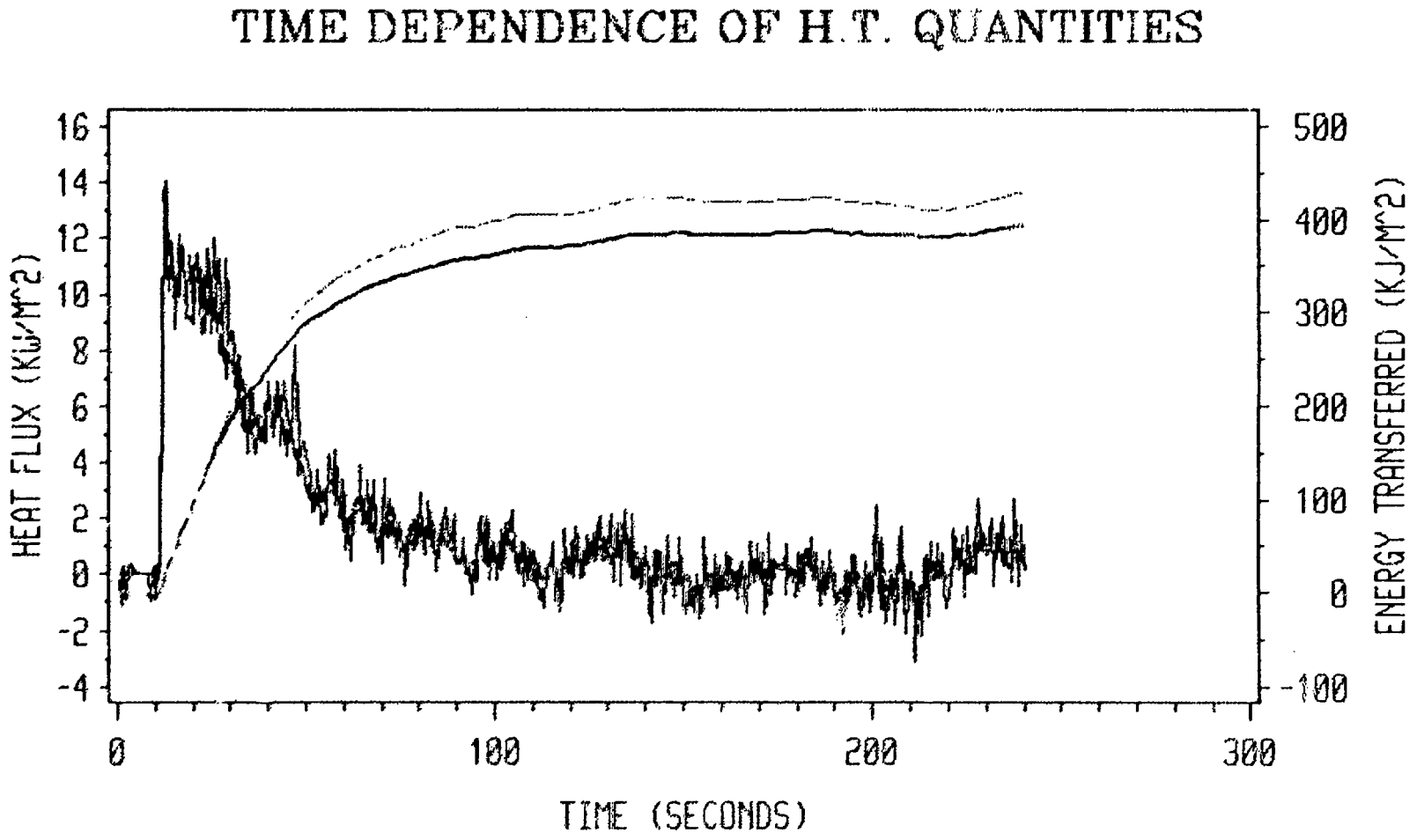
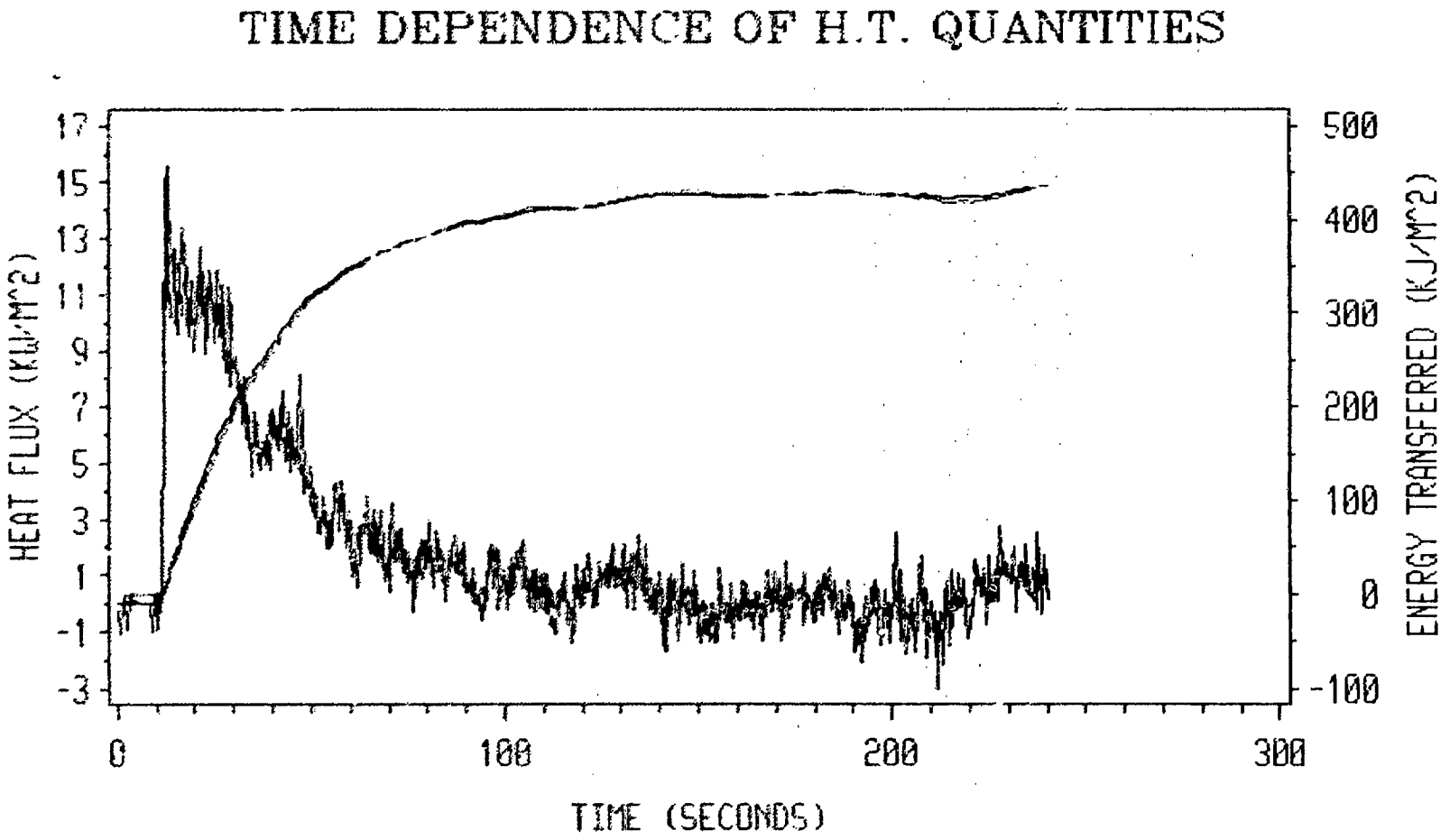


FIG. 5 COMPARISON PLOTS FOR HEAT FLUX TRANS. OUTPUT AND T.C. ISOLATED PAIR. FOR DATA AT 0.5 SEC. INTERVALS. W/ MULT. FACTOR OF 1.000. RAW DATA--FN 43



EXP. HEAT FLUX
TOTAL EXP. ENERGY
NUM. HEAT FLUX
TOTAL NUM. ENERGY

FIG. 6 COMPARISON PLOTS FOR HEAT FLUX TRANS. OUTPUT AND T.C. ISOLATED PAIR. FOR DATA AT 0.5 SEC. INTERVALS. W/ MULT. FACTOR OF 1.105. RAW DATA--FN 43

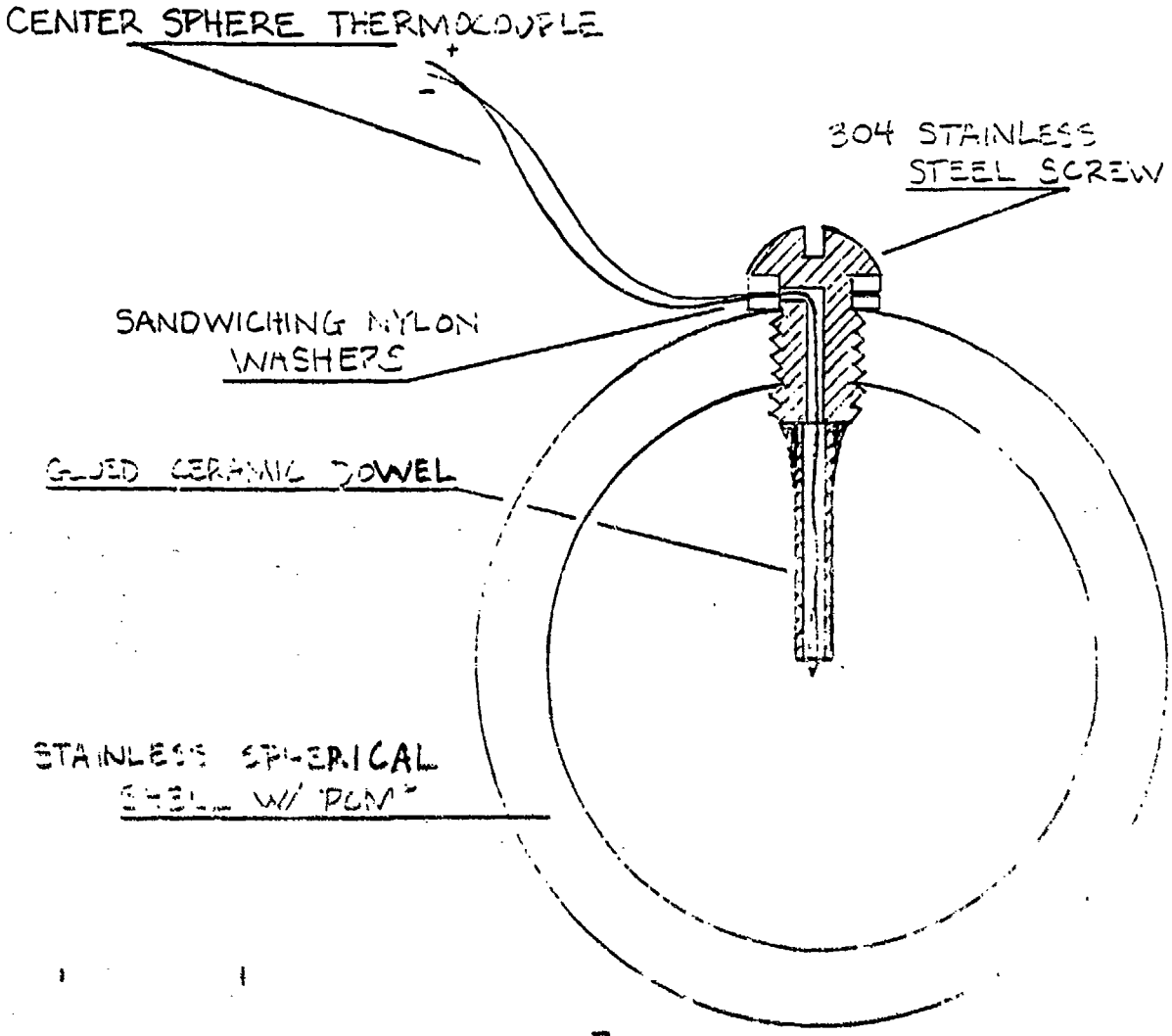


FIGURE 2 DETAIL OF CENTER THERMOCOUPLE PROBE.

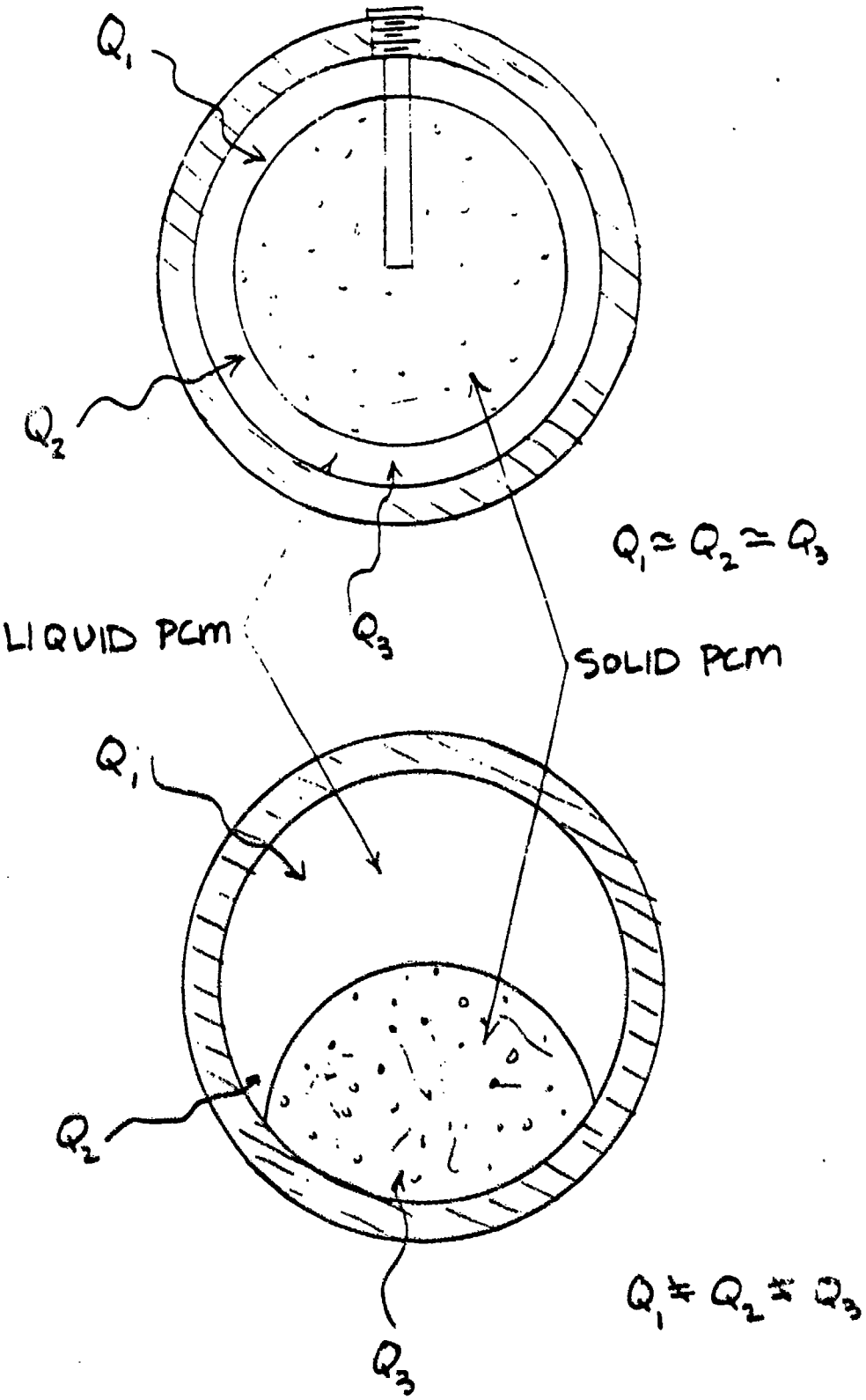
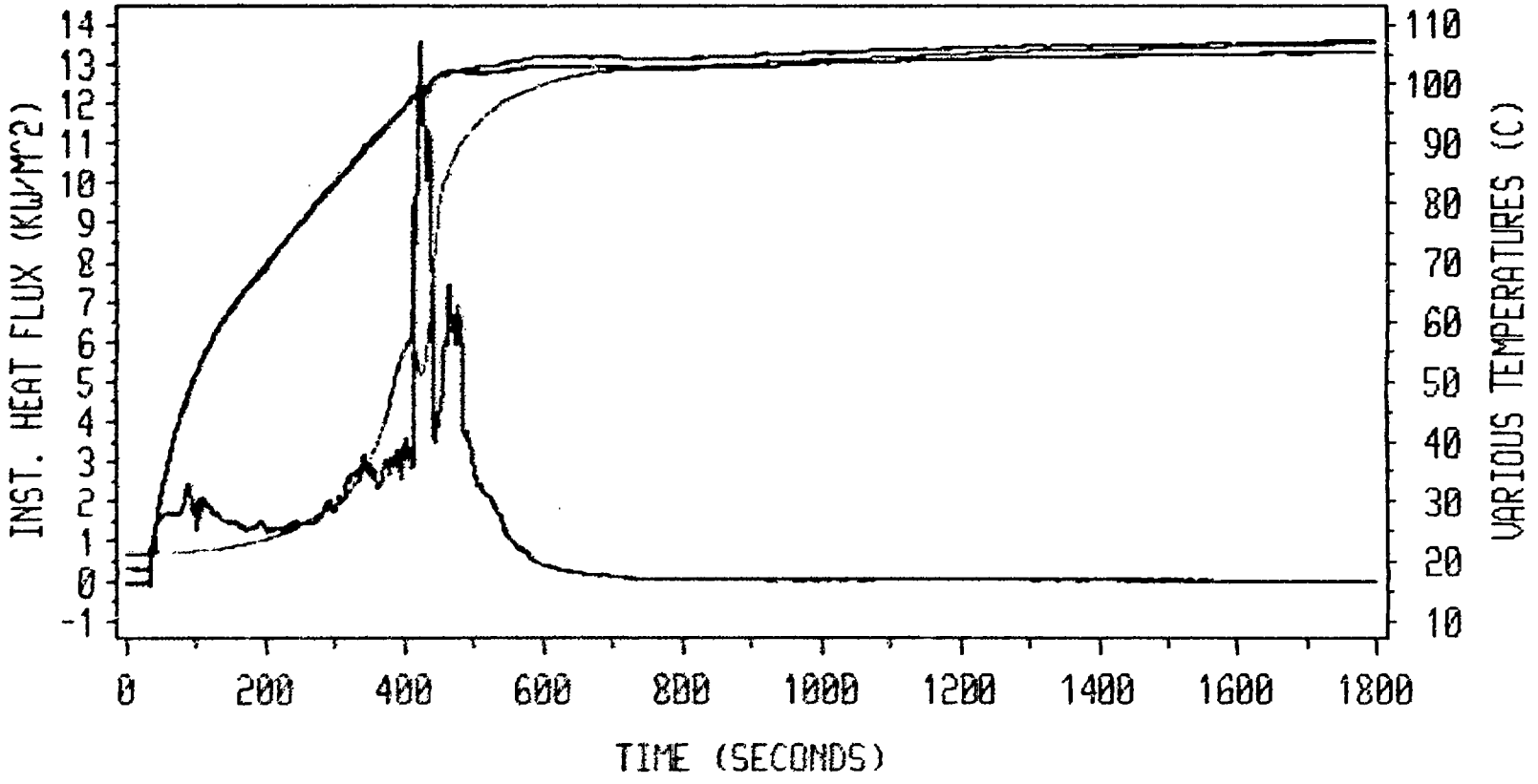


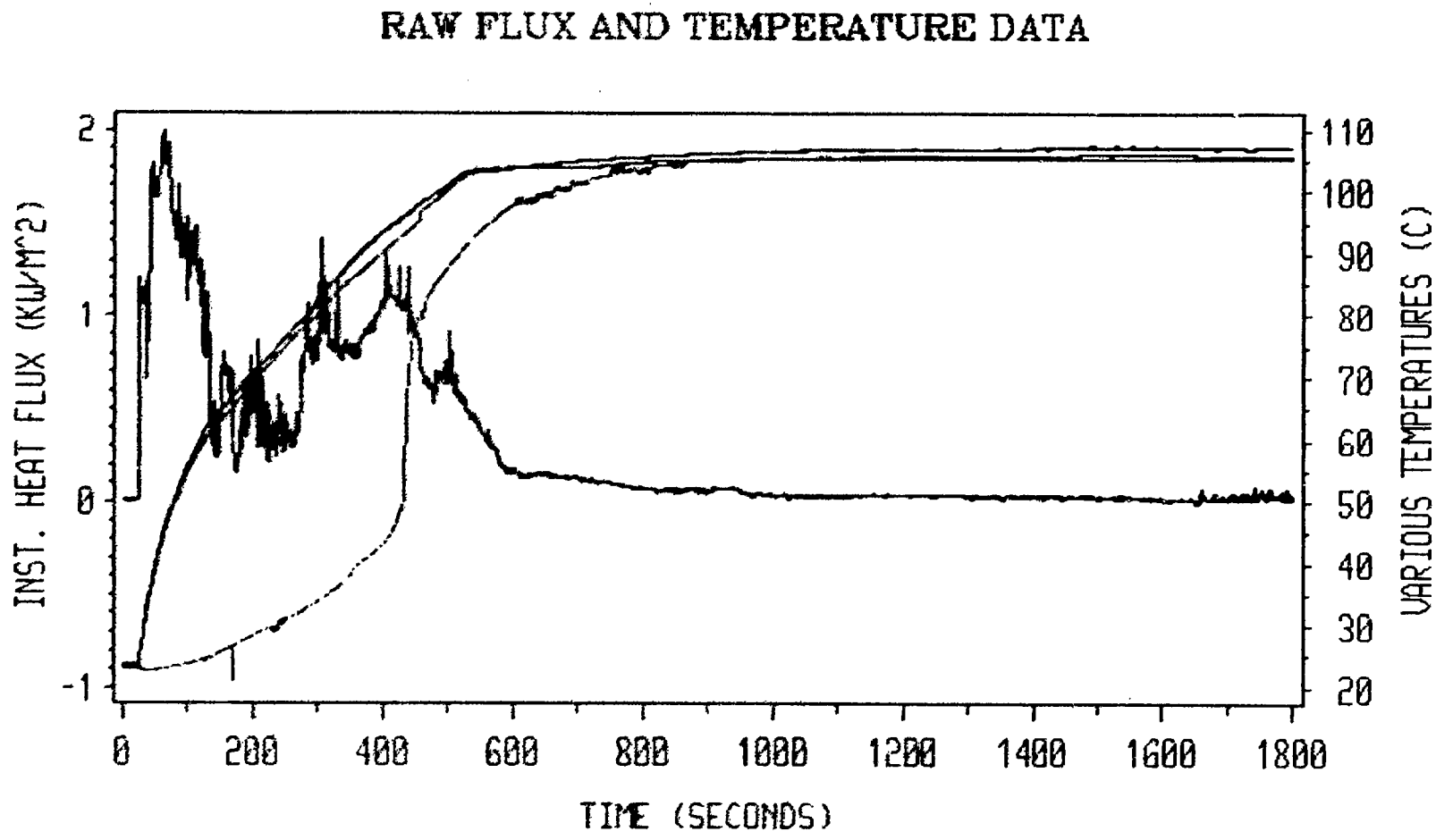
FIGURE 8 FLUX DISTRIBUTION FOR MELT PROCESS INSIDE SPHERE

RAW FLUX AND TEMPERATURE DATA



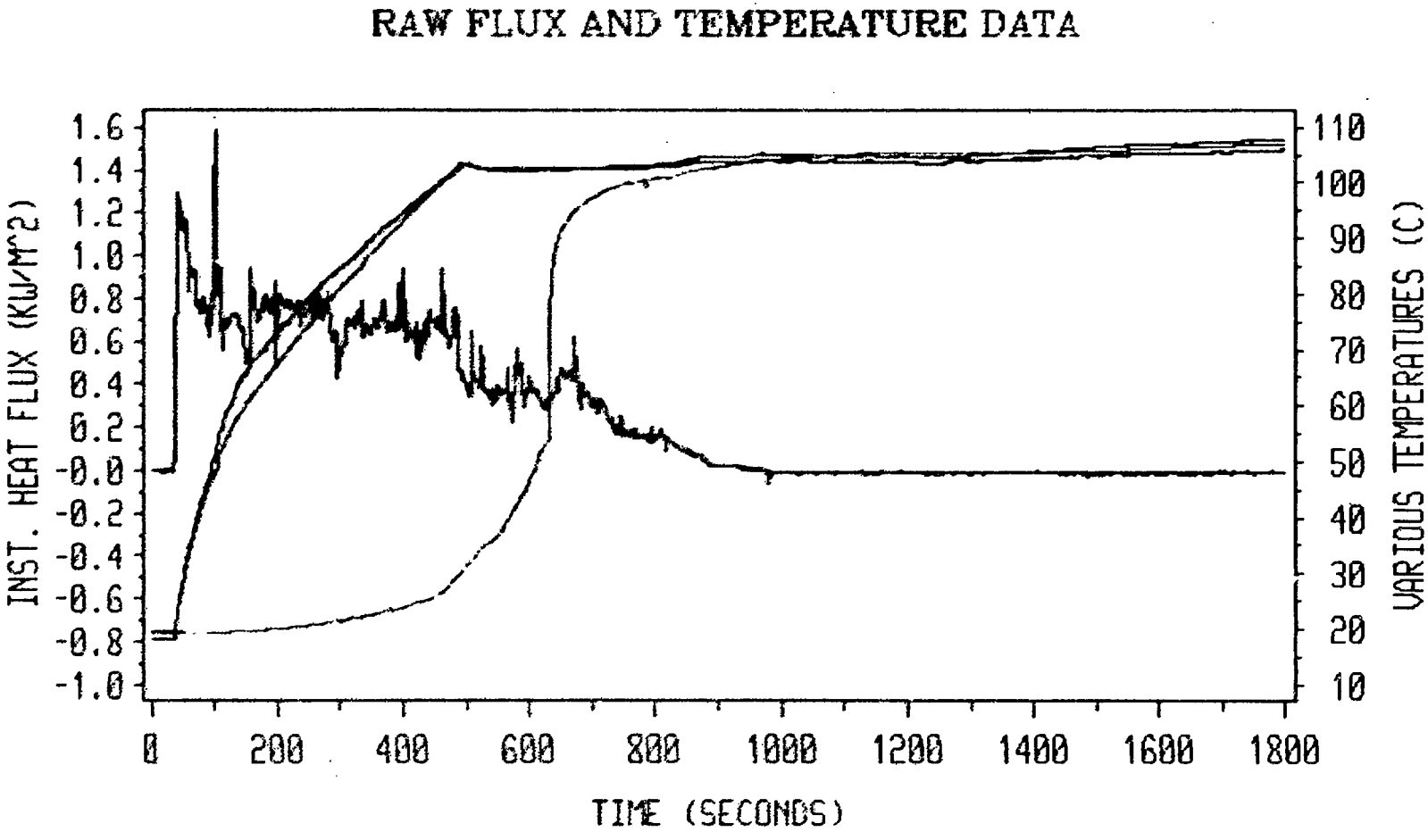
HEAT FLUX TRANSDUCER
H.F. TRANSDUCER T.C.
CENTER SPHERE T.C.
AMBIENT TEMPERATURE

FIG. 9 CMS FN-59. PLOTS OF TEMPERATURE AND HEAT FLUX DATA WITH TIME FOR 2.0" DIAMETER ISOLATED SPHERE. DATA AT 1 SEC. MC5600 FN-MAIN204.



HEAT FLUX TRANSDUCER
H.F. TRANSDUCER T.C.
CENTER SPHERE T.C.
AMBIENT TEMPERATURE

FIG. 10 CMS FN-58. PLOTS OF TEMPERATURE AND HEAT FLUX DATA WITH TIME FOR 2.0" DIAMETER ARRAYED SPHERE. DATA AT 1 SEC. MC5600 FN-MAIN203.



HEAT FLUX TRANSDUCER
H.F. TRANSDUCER T.C.
CENTER SPHERE T.C.
AMBIENT TEMPERATURE

FIG. 11 CMS FN-56. PLOTS OF TEMPERATURE AND HEAT FLUX DATA WITH TIME FOR 2.5" DIAMETER ISOLATED SPHERE. DATA AT 1 SEC. MC5600 FN-MAIN252.

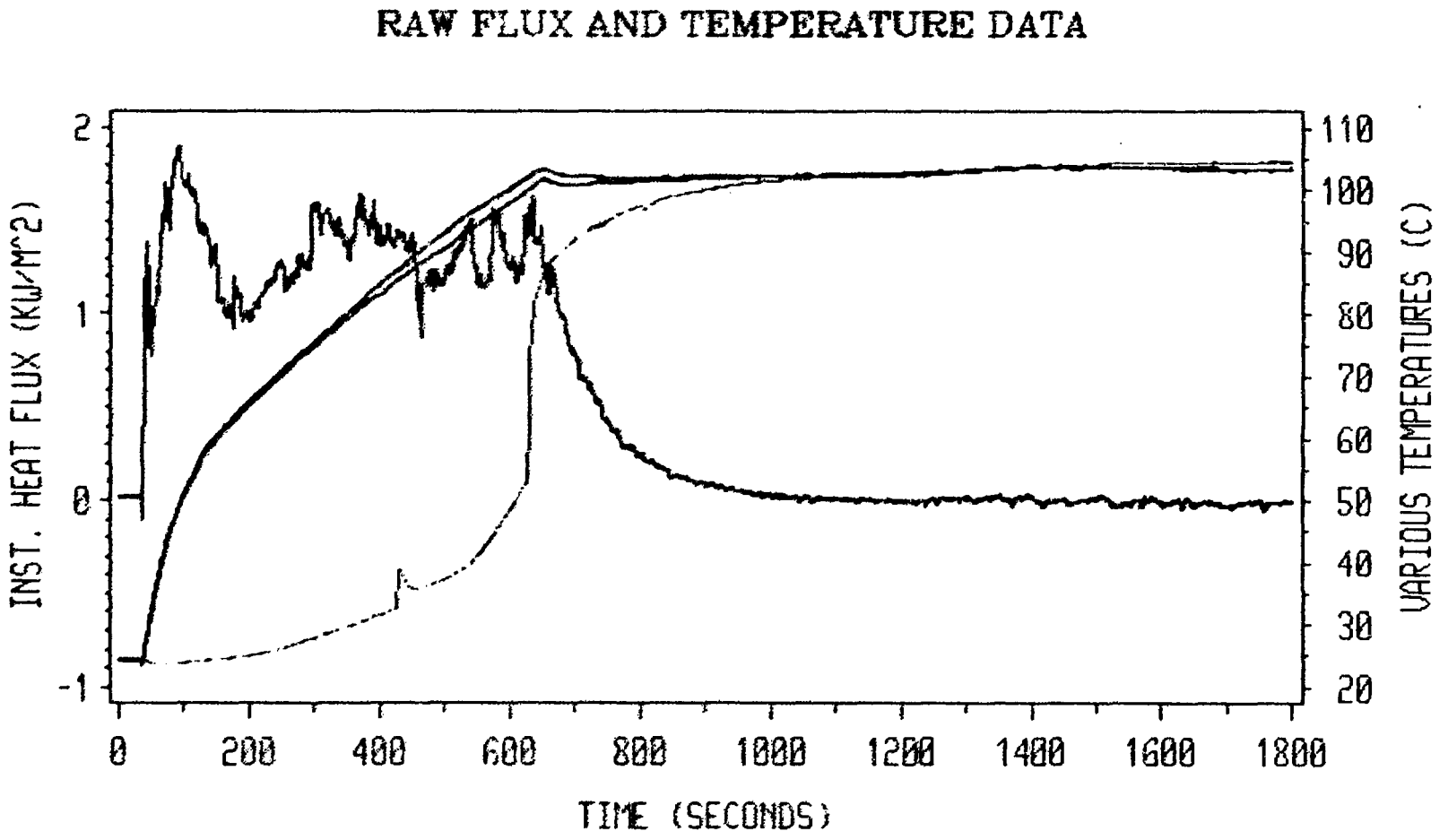


FIG. 12 CMS FN-57. PLOTS OF TEMPERATURE AND HEAT FLUX DATA WITH TIME FOR 2.5" DIAMETER ARRAYED SPHERE. DATA AT 1 SEC. MC5600 FN-MAIN253.

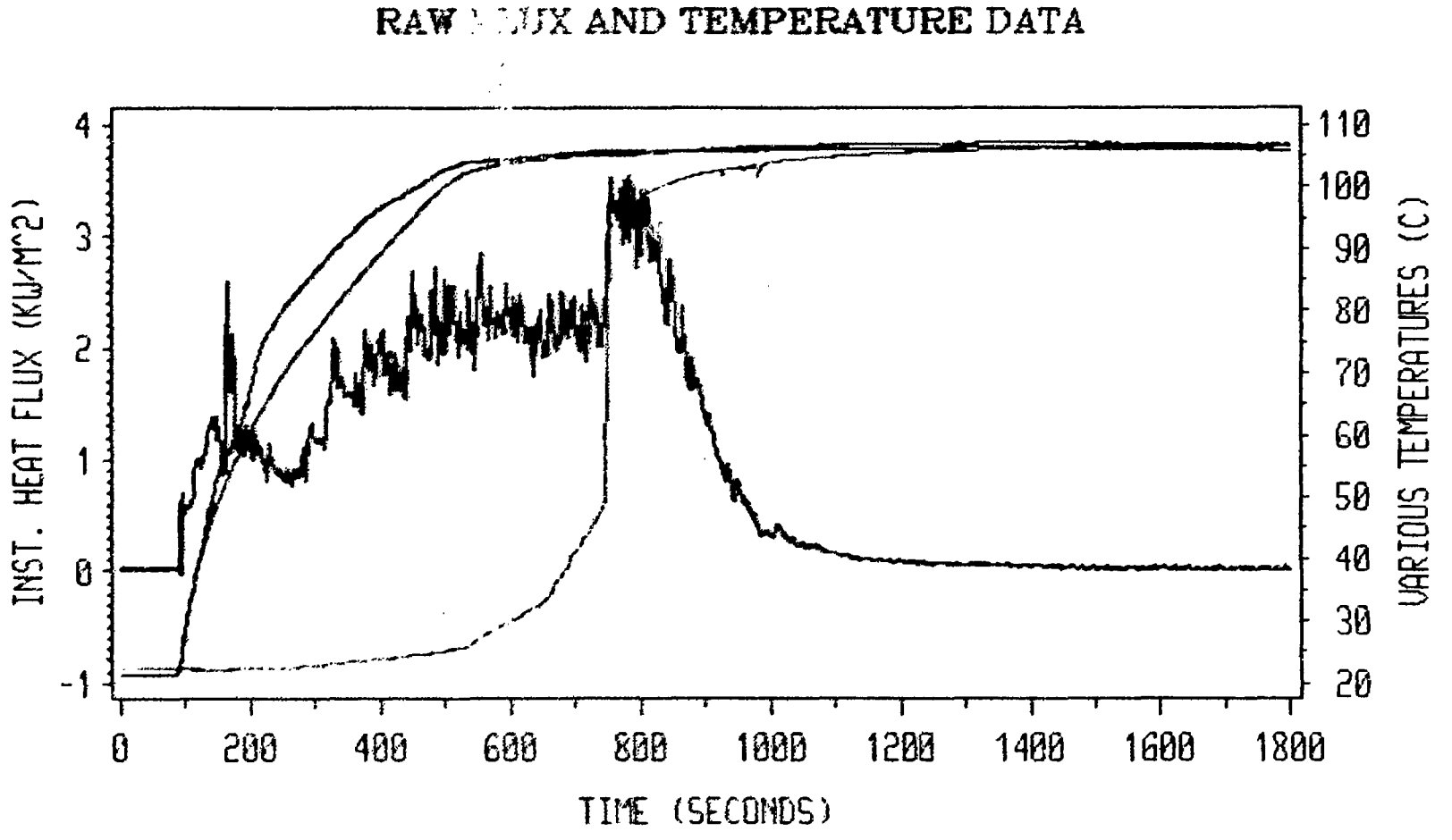
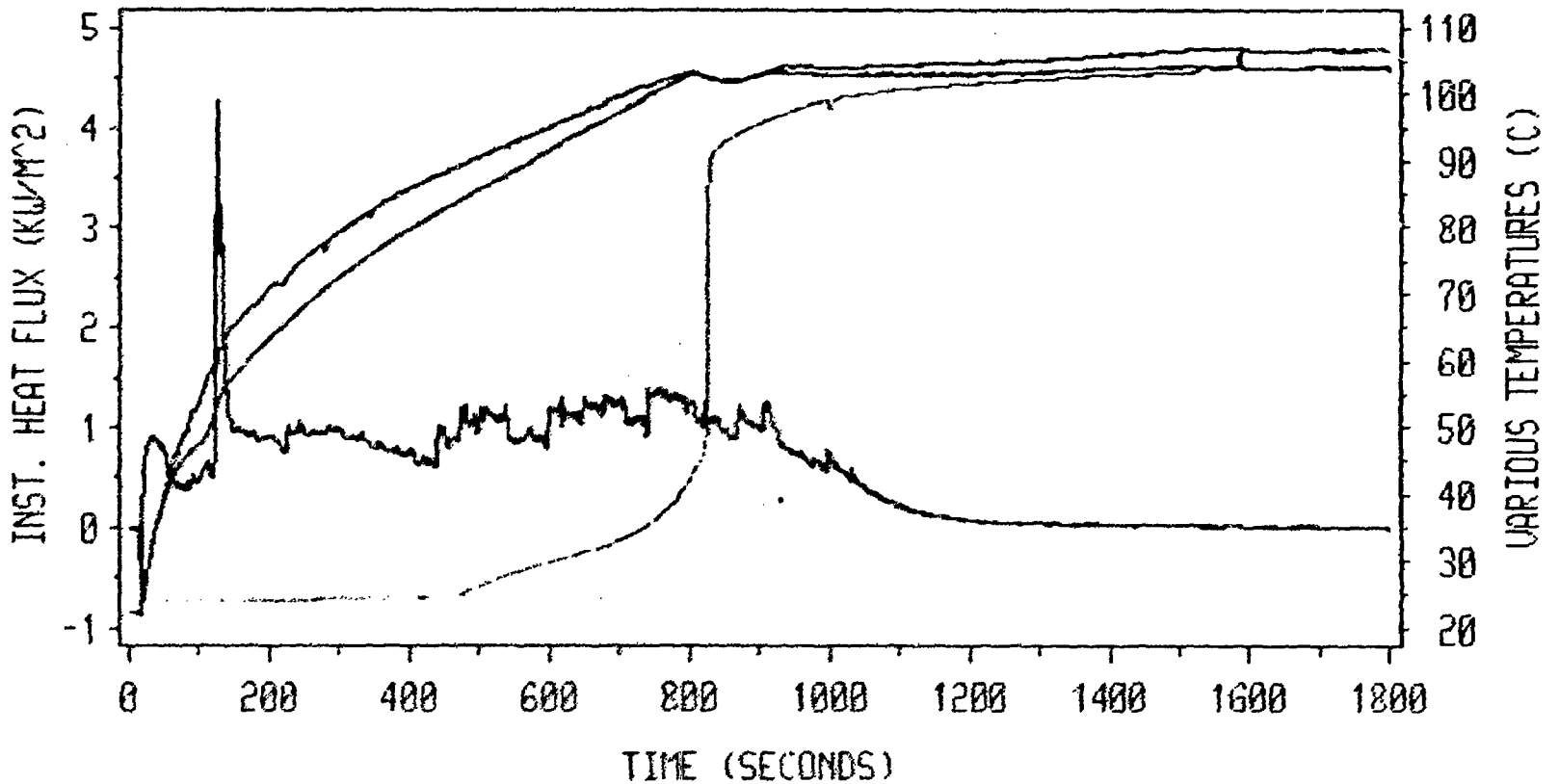


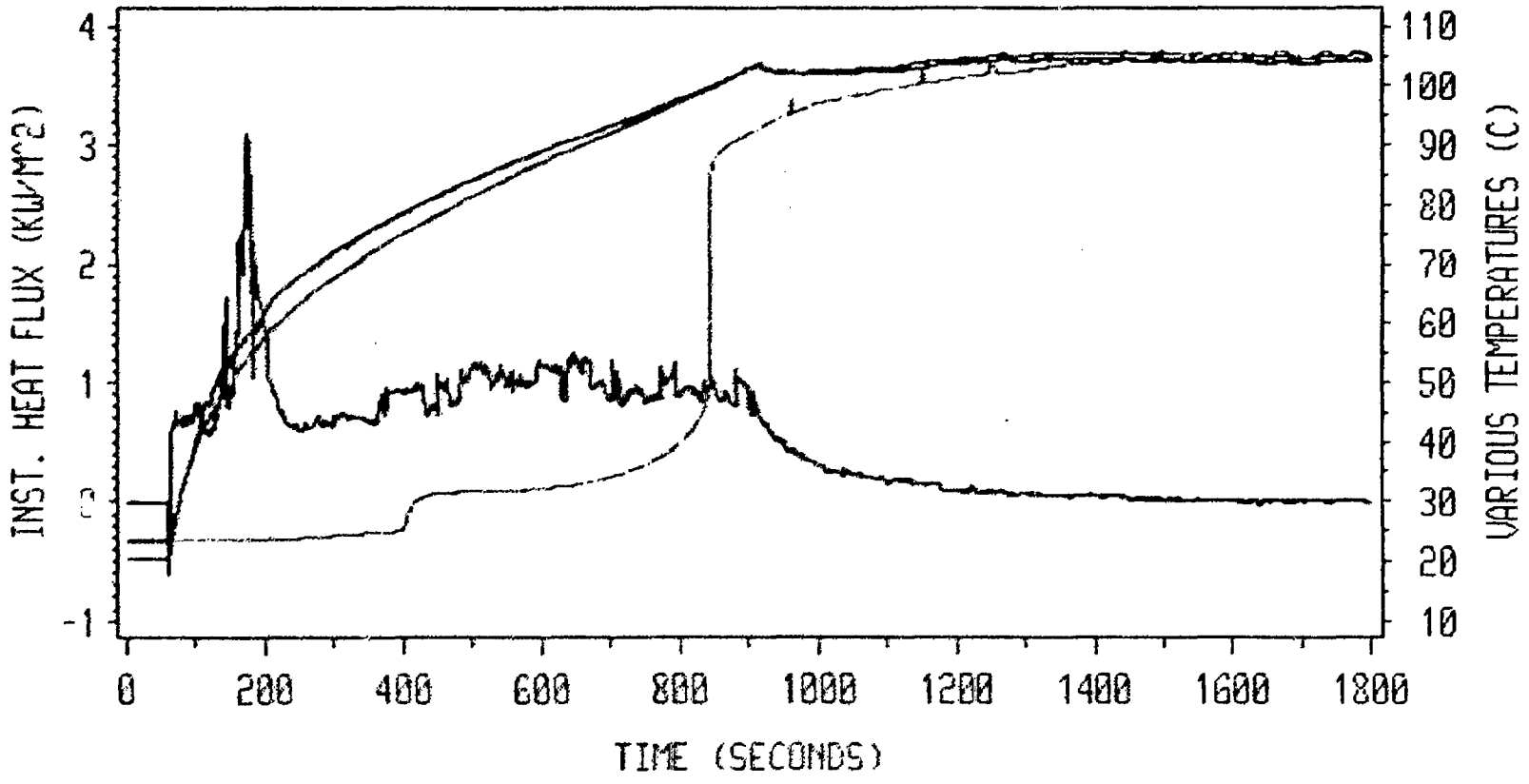
FIG. 13 CMS FN-51. PLOTS OF TEMPERATURE AND HEAT FLUX DATA WITH TIME FOR 3.0" DIAMETER ISOLATED SPHERE. DATA AT 1 SEC. MC5600 FN-MAIN301.

RAW FLUX AND TEMPERATURE DATA



HEAT FLUX TRANSDUCER
H.F. TRANSDUCER T.C.
CENTER SPHERE T.C.
AMBIENT TEMPERATURE

FIG. 4 CMS FN-53. PLOTS OF TEMPERATURE AND HEAT FLUX DATA WITH TIME FOR 3.0" DIAMETER ARRAYED SPHERE. DATA AT 1 SEC. MC5600 FN-MAIN303.



HEAT FLUX TRANSDUCER
H.F. TRANSDUCER T.C.
CENTER SPHERE T.C.
AMBIENT TEMPERATURE

FIG. 15 CMS FN-52. PLOTS OF TEMPERATURE AND HEAT FLUX DATA WITH TIME FOR 3.0" DIAMETER ARRAYED SPHERE. DATA AT 1 SEC. MC5600 FN-MAIN302.

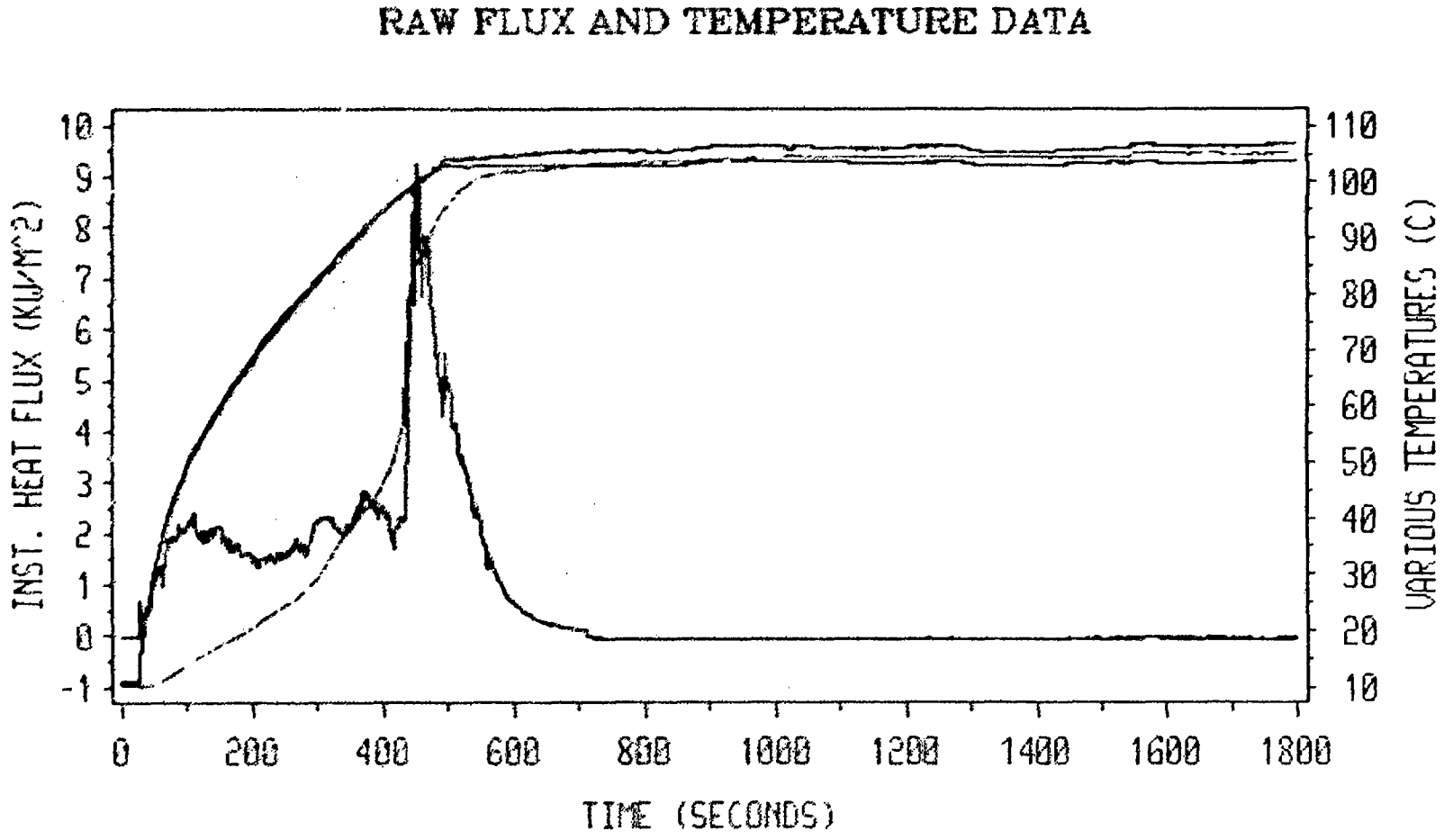


FIG. 16 CMS FN-60. PLOTS OF TEMPERATURE AND HEAT FLUX DATA WITH TIME FOR 2.0" DIAMETER ISOLATED SPHERE. DATA AT 1 SEC. MC5680 FN-MAIN205.

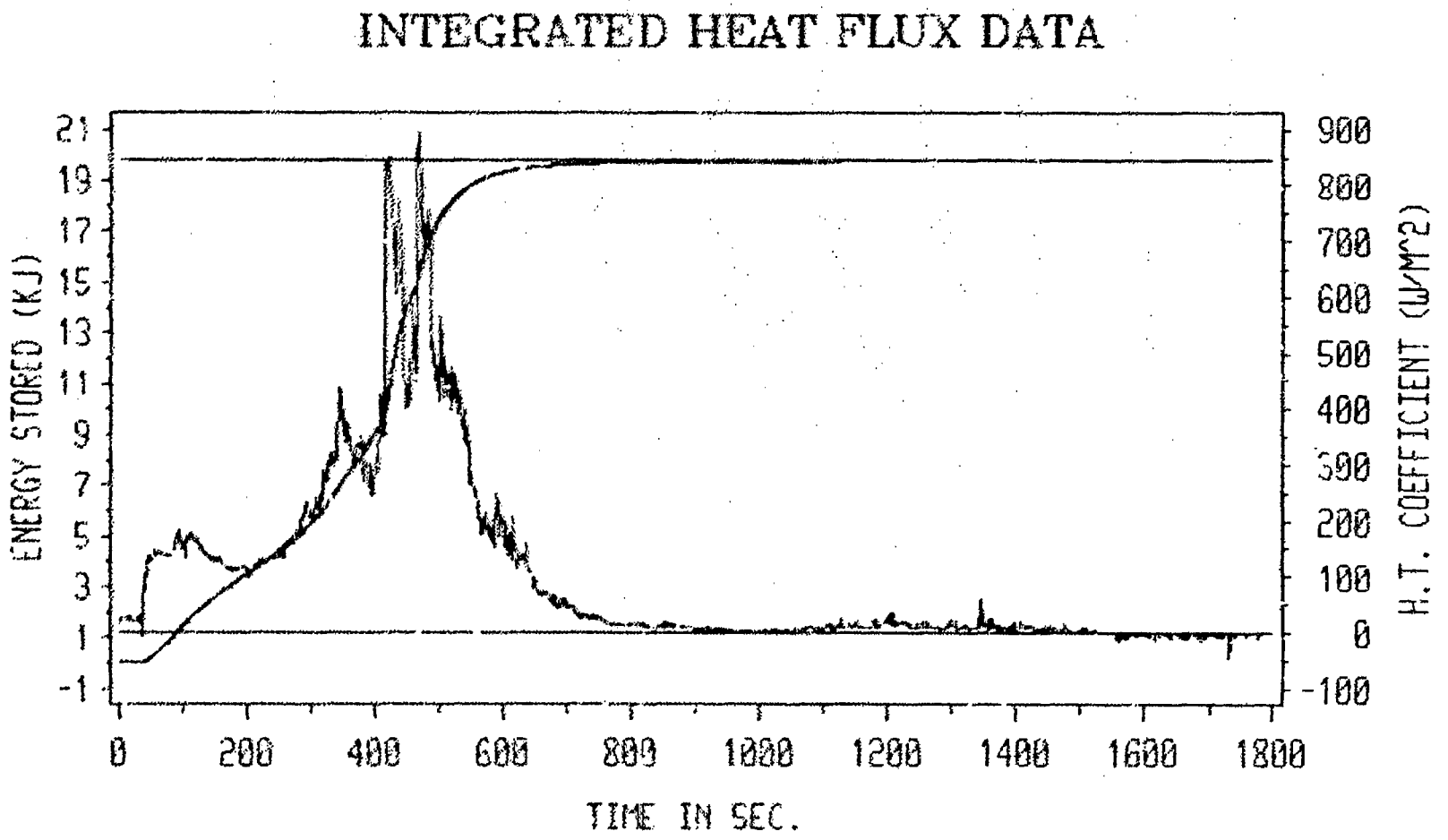


FIG. 17 FN-59. PLOT OF EXPERIMENTAL ENERGY AND H.T.C.
FOR 2.0" DIAM. ISOLATED SPHERE WITH MF=1.646.
FOR SAMPLING AT 1.0 SECOND; MC5600 FN-MAIN204.
TEMPERATURE DIFF. IN H CALC. USING WALL TEMP.

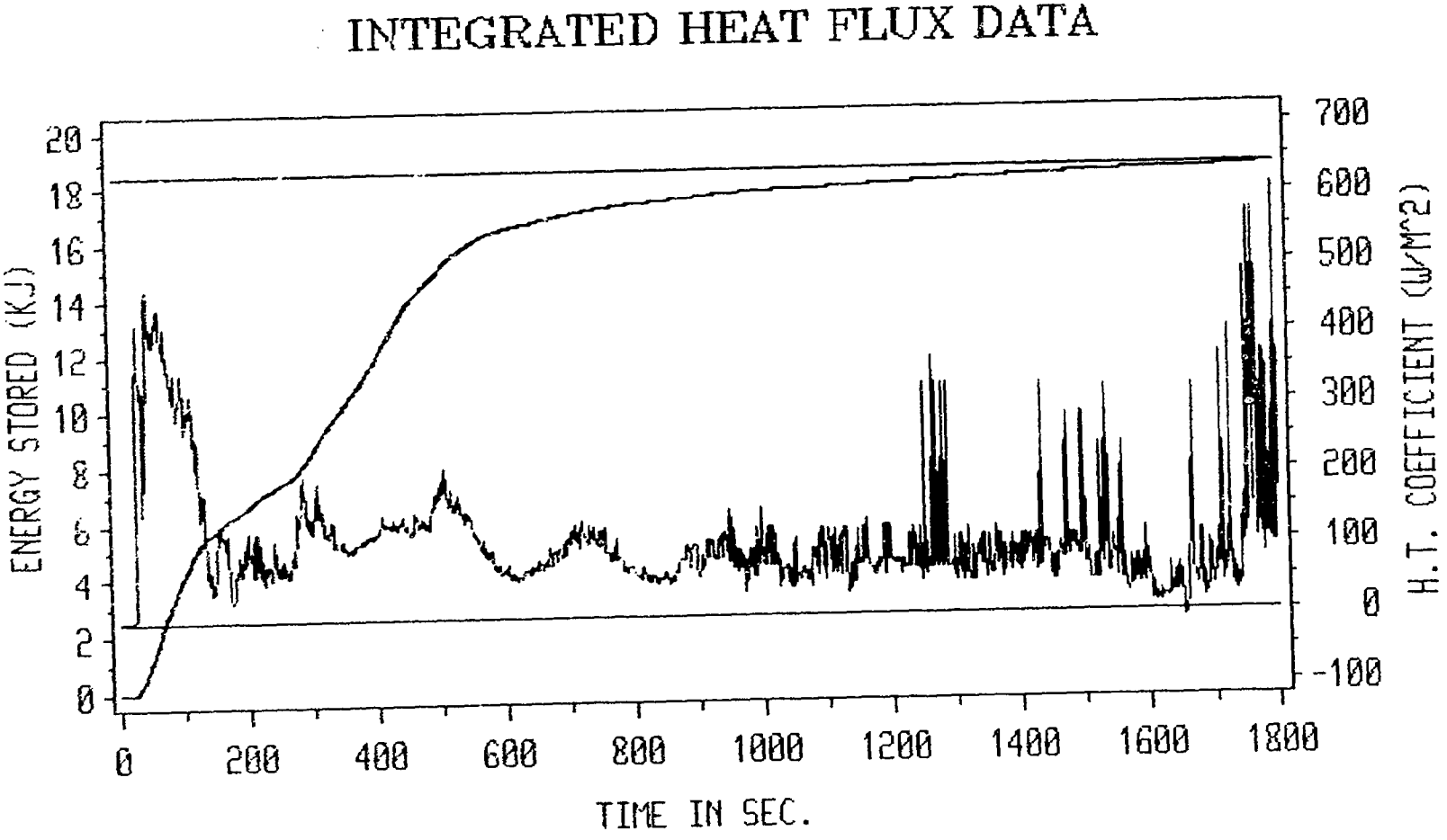
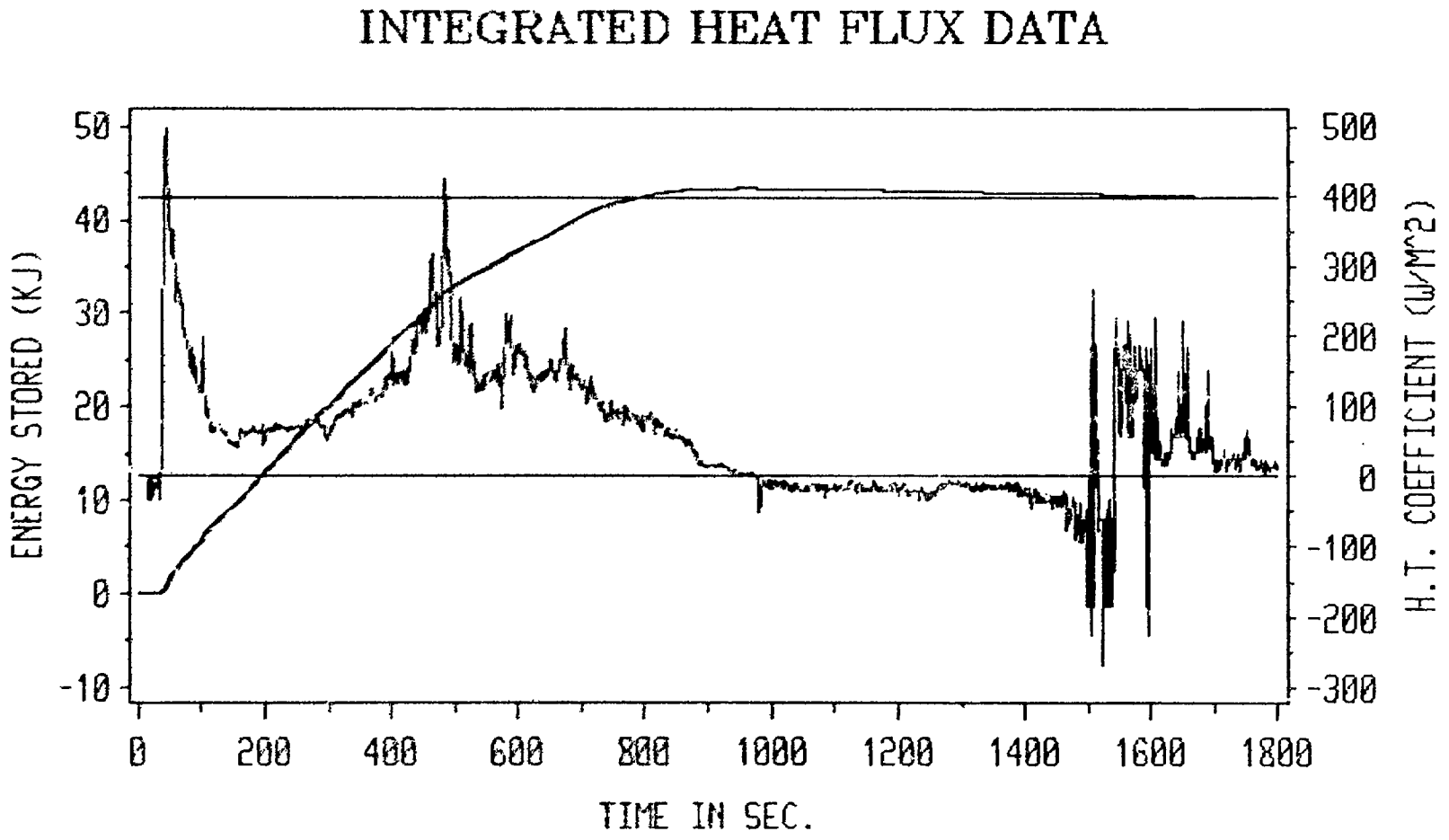


FIG. 18 FN-58. PLOT OF EXPERIMENTAL ENERGY AND H.T.C. FOR 2.0' DIAM. ARRAYED SPHERE WITH MF=4.448. FOR SAMPLING AT 1.0 SECOND; MC5600 FN-MAIN203. TEMPERATURE DIFF. IN H CALC. USING WALL TEMP.



TRANSD. ENERGY STORAGE
EXPECTED ENERGY STORAGE
HEAT TRANS. COEFF.
ZERO REFERENCE H.T. COEFF.

FIG. 9 FN-56. PLOT OF EXPERIMENTAL ENERGY AND H.T.C.
FOR 2.5" DIAM. ISOLATED SPHERE WITH MF=7.767.
FOR SAMPLING AT 1.0 SECOND; MC5600 FN-MAIN252.
TEMPERATURE DIFF. IN H CALC. USING WALL TEMP.

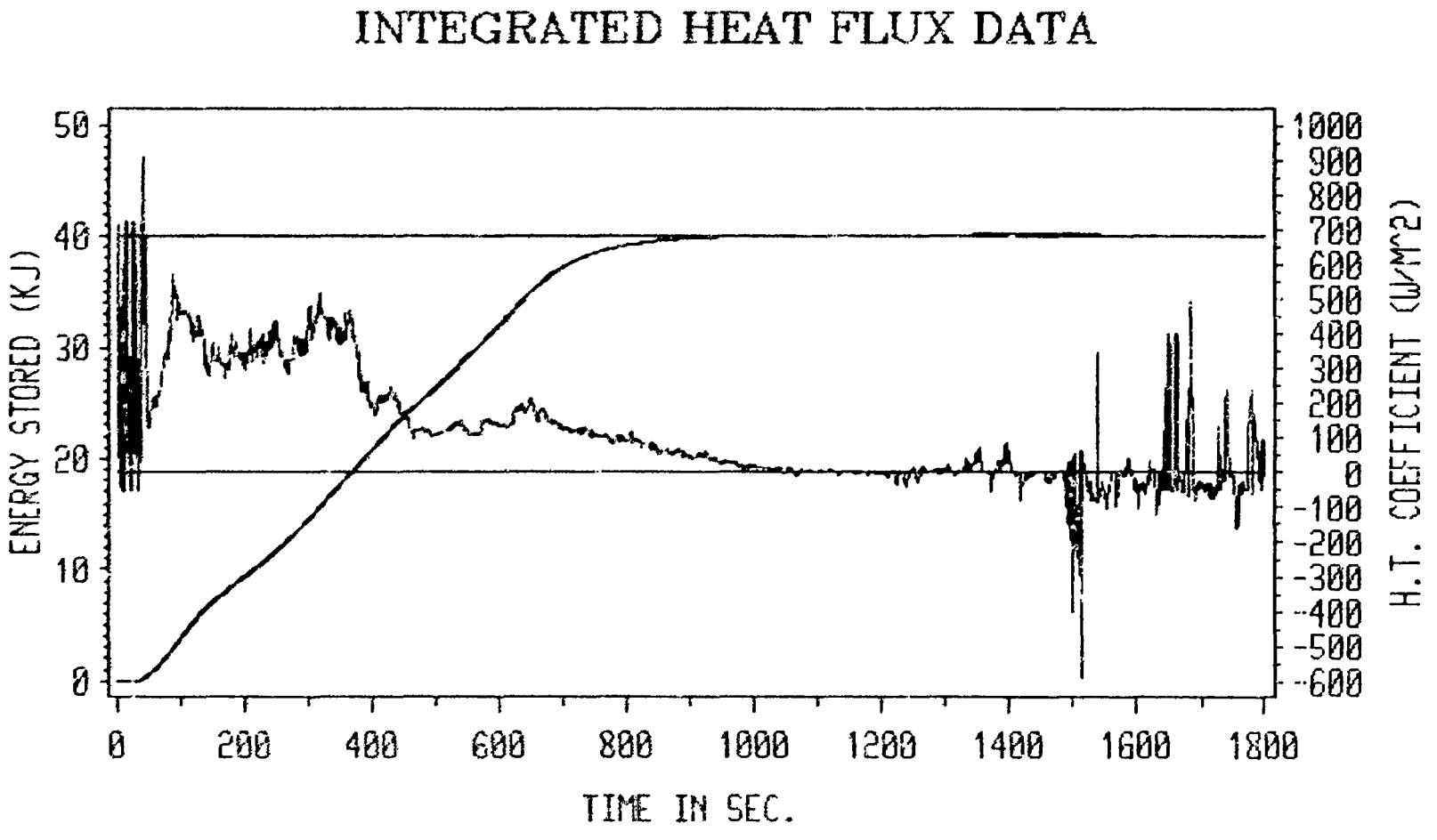


FIG 20 FN-57. PLOT OF EXPERIMENTAL ENERGY AND H.T.C.
FOR 2.5' DIAM. ARRAYED SPHERE WITH MF=3.390.
FOR SAMPLING AT 1.0 SECOND; MC5600 FN-MAIN253.
TEMPERATURE DIFF. IN H CALC. USING WALL TEMP.

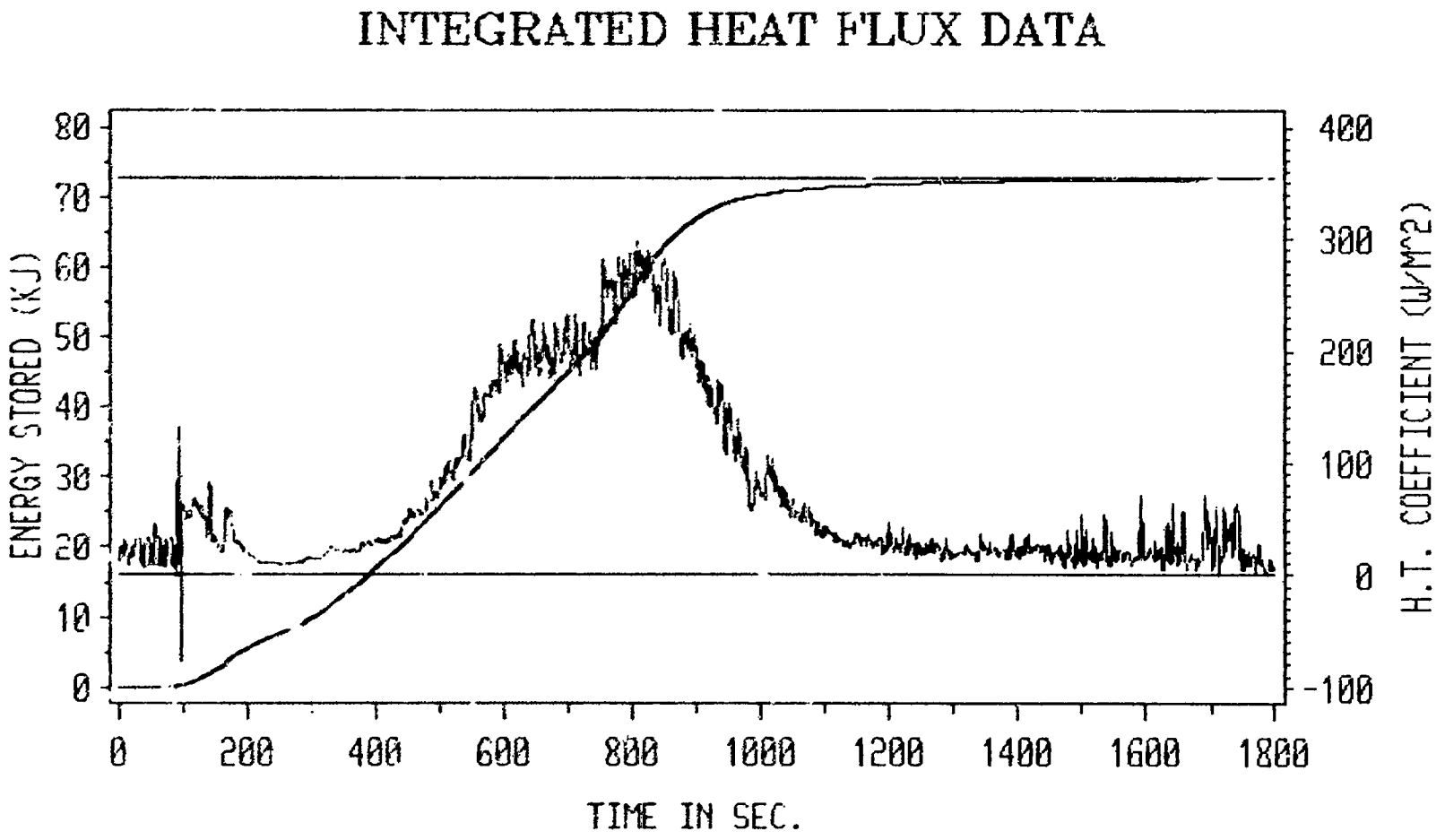
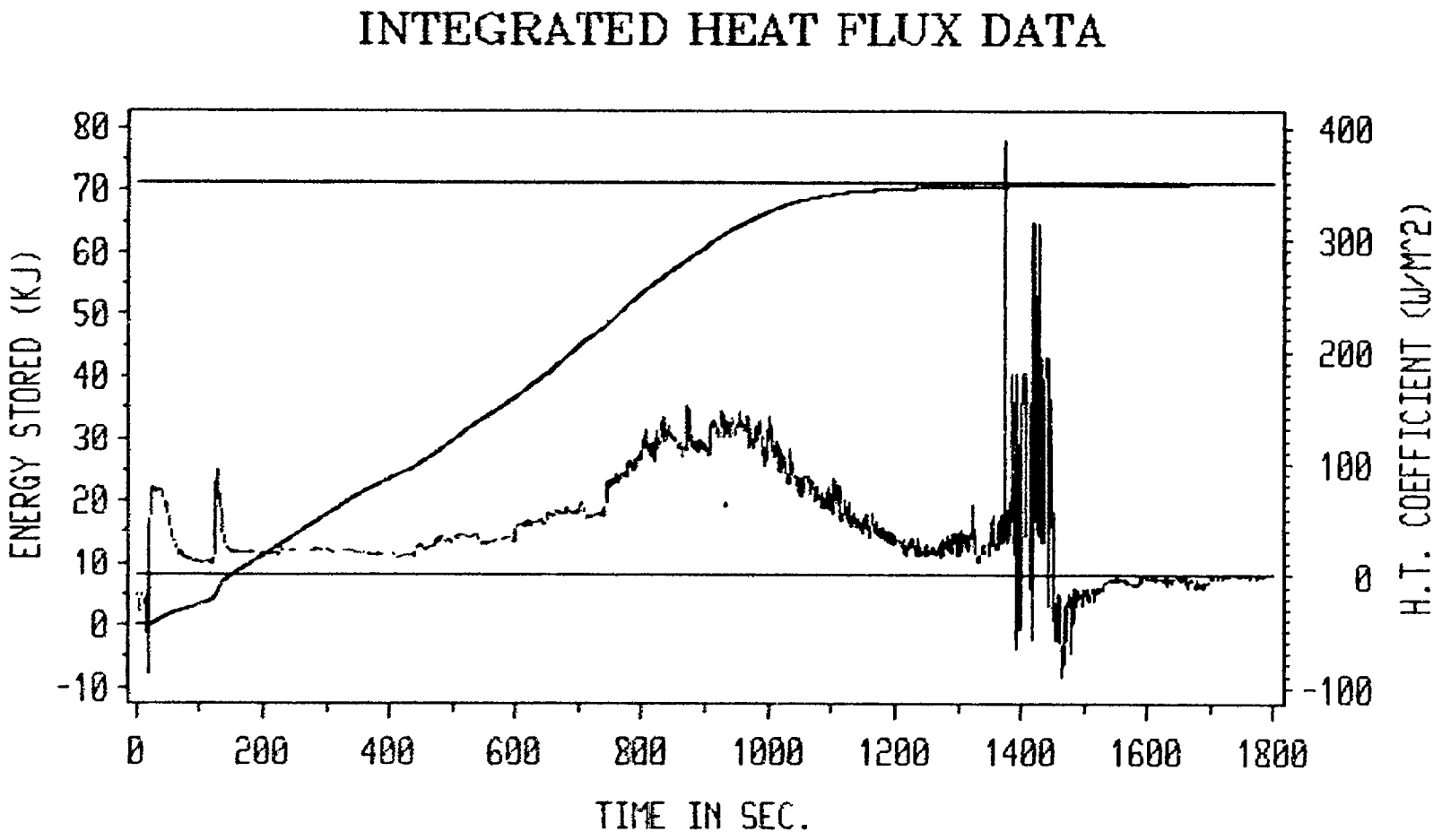
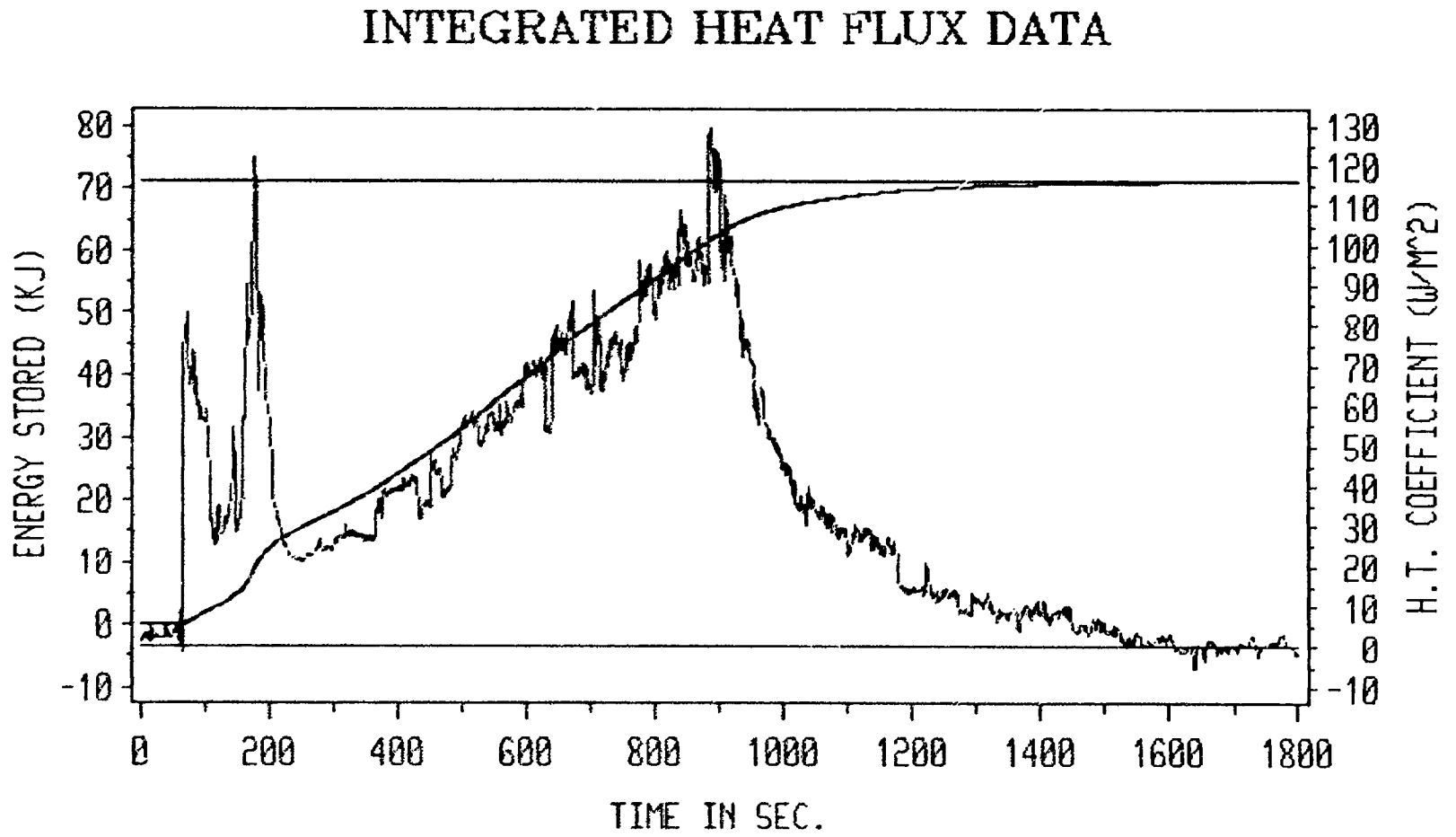


FIG 21 FN-51. PLOT OF EXPERIMENTAL ENERGY AND H.T.C.
FOR 3.0" DIAM. ISOLATED SPHERE WITH MF=2.387.
FOR SAMPLING AT 1.0 SECOND; MC5600 FN-MAIN301.
TEMPERATURE DIFF. IN H CALC. USING WALL TEMP.



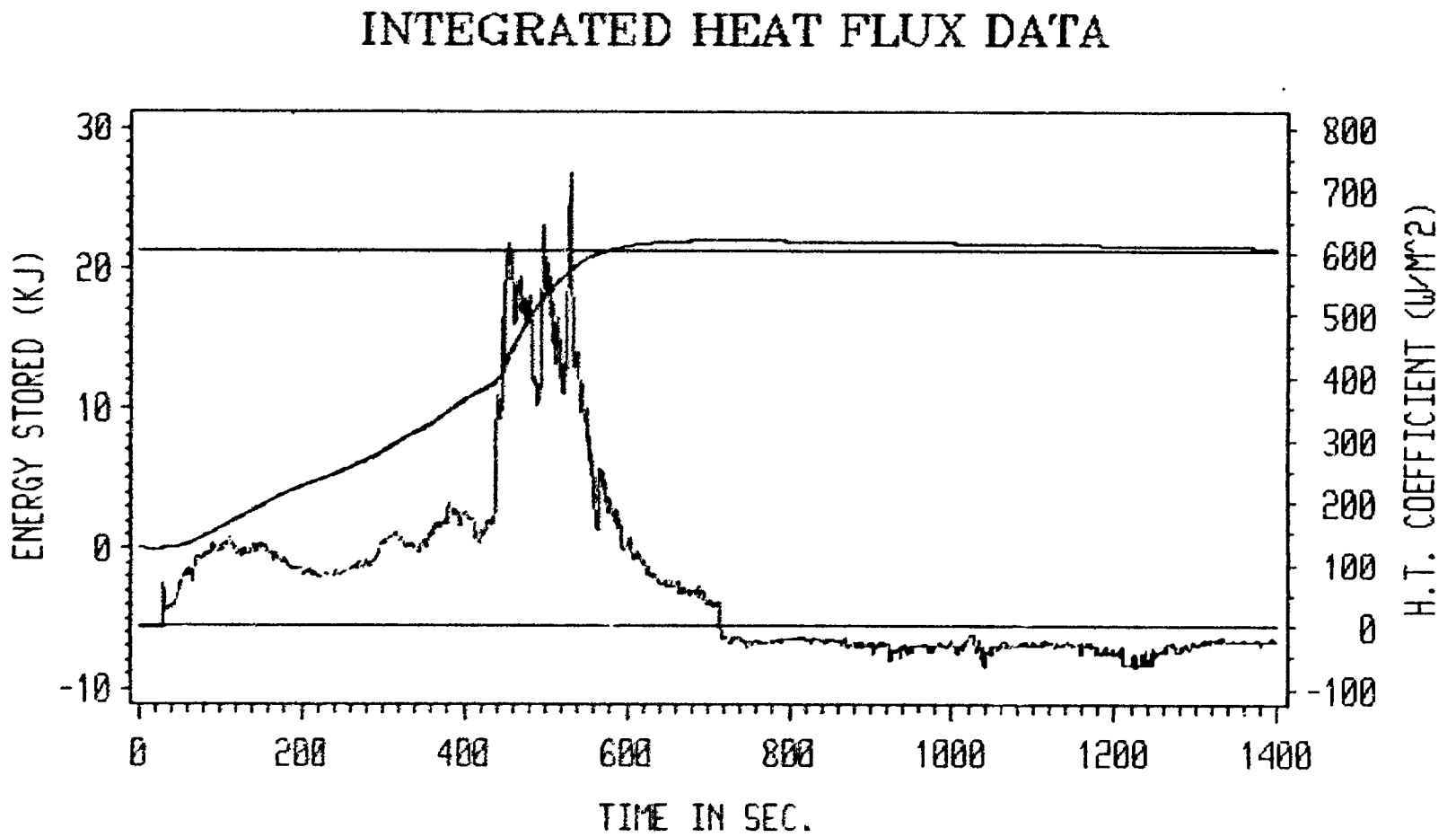
TRANSD. ENERGY STORAGE
EXPECTED ENERGY STORAGE
HEAT TRANS. COEFF.
ZERO REFERENCE H.T. COEFF.

FIG 22 FN-53. PLOT OF EXPERIMENTAL ENERGY AND H.T.C.
FOR 3.0' DIAM. ARRAYED SPHERE WITH MF=3.837.
FOR SAMPLING AT 1.0 SECOND; MC5600 FN-MAIN303.
TEMPERATURE DIFF. IN H CALC. USING WALL TEMP.



TRANSD. ENERGY STORAGE
EXPECTED ENERGY STORAGE
HEAT TRANS. COEFF.
ZERO REFERENCE H.T. COEFF.

FIG.23 FN-52. PLOT OF EXPERIMENTAL ENERGY AND H.T.C.
FOR 3.0' DIAM. ARRAYED SPHERE WITH MF=4.336.
FOR SAMPLING AT 1.0 SECOND; MC5600 FN-MAIN302.
TEMPERATURE DIFF. IN H CALC. USING WALL TEMP.



TRANSD. ENERGY STORAGE
EXPECTED ENERGY STORAGE
HEAT TRANS. COEFF.
ZERO REFERENCE H.T. COEFF.

FIG. 24 FN-60. PLOT OF EXPERIMENTAL ENERGY AND H.T.C.
FOR 2.0" DIAM. ISOLATED SPHERE WITH MF=1.88.
FOR SAMPLING AT 1.0 SECOND; MC5600 FN-MAIN205.
TEMPERATURE DIFF. IN H CALC. USING WALL TEMP.

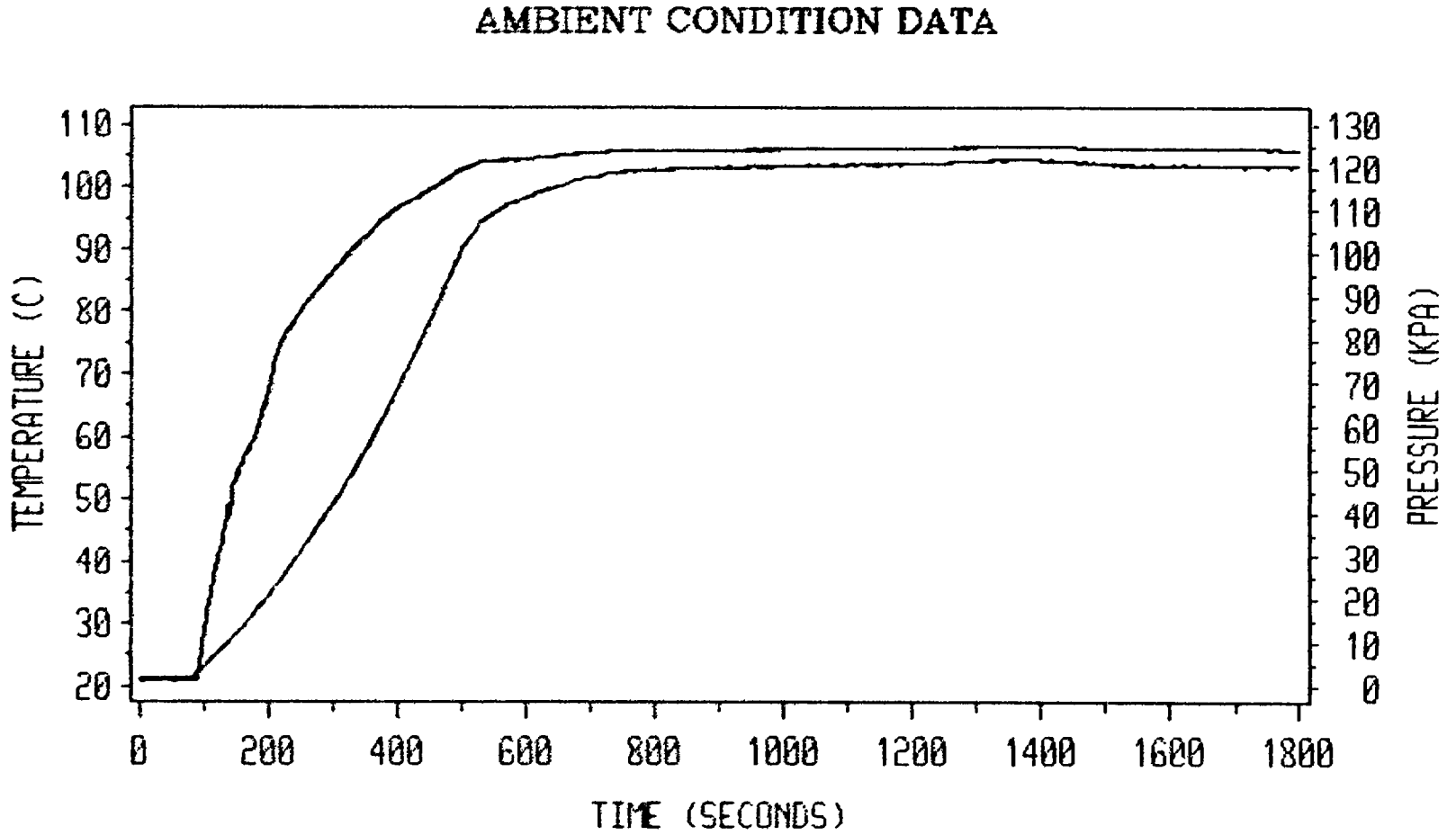
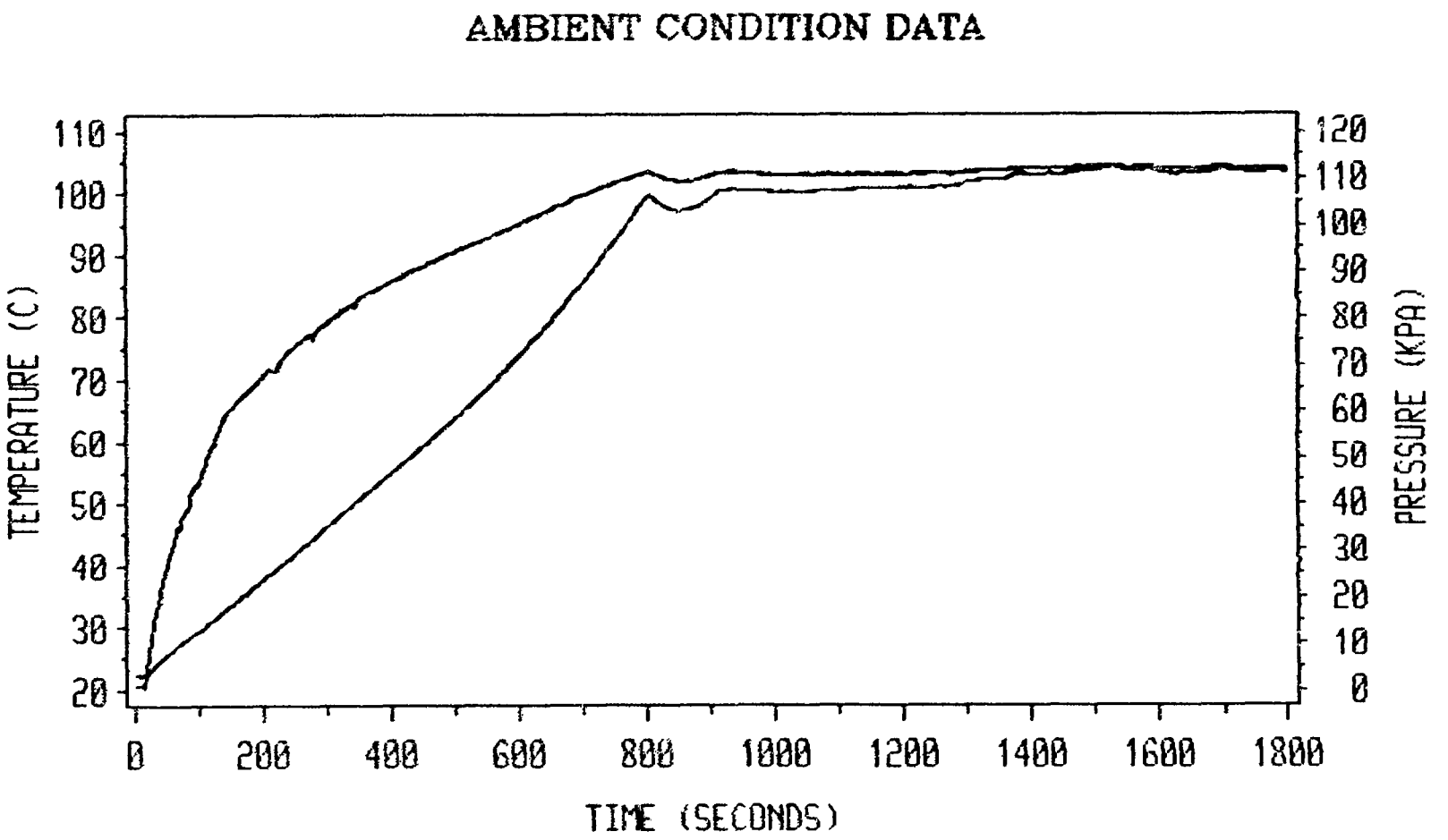


FIG. 25 CMS FN-51. PLOTS OF TEMPERATURE AND PRESSURE IN SPHERE AMBIENCE AS A FUNCTION OF TIME FOR 3.0" DIAMETER ISOLATED SPHERE DATA AT 1 SEC. MC5600 FN-51.



AMBIENT TEMPERATURE
AMBIENT PRESSURE

FIG. 26 CMS FN-53. PLOTS OF TEMPERATURE AND PRESSURE IN SPHERE AMBIENCE AS A FUNCTION OF TIME FOR 3.0" DIAMETER ARRAYED SPHERE DATA AT 1 SEC. MC5600 FN-MAIN303.

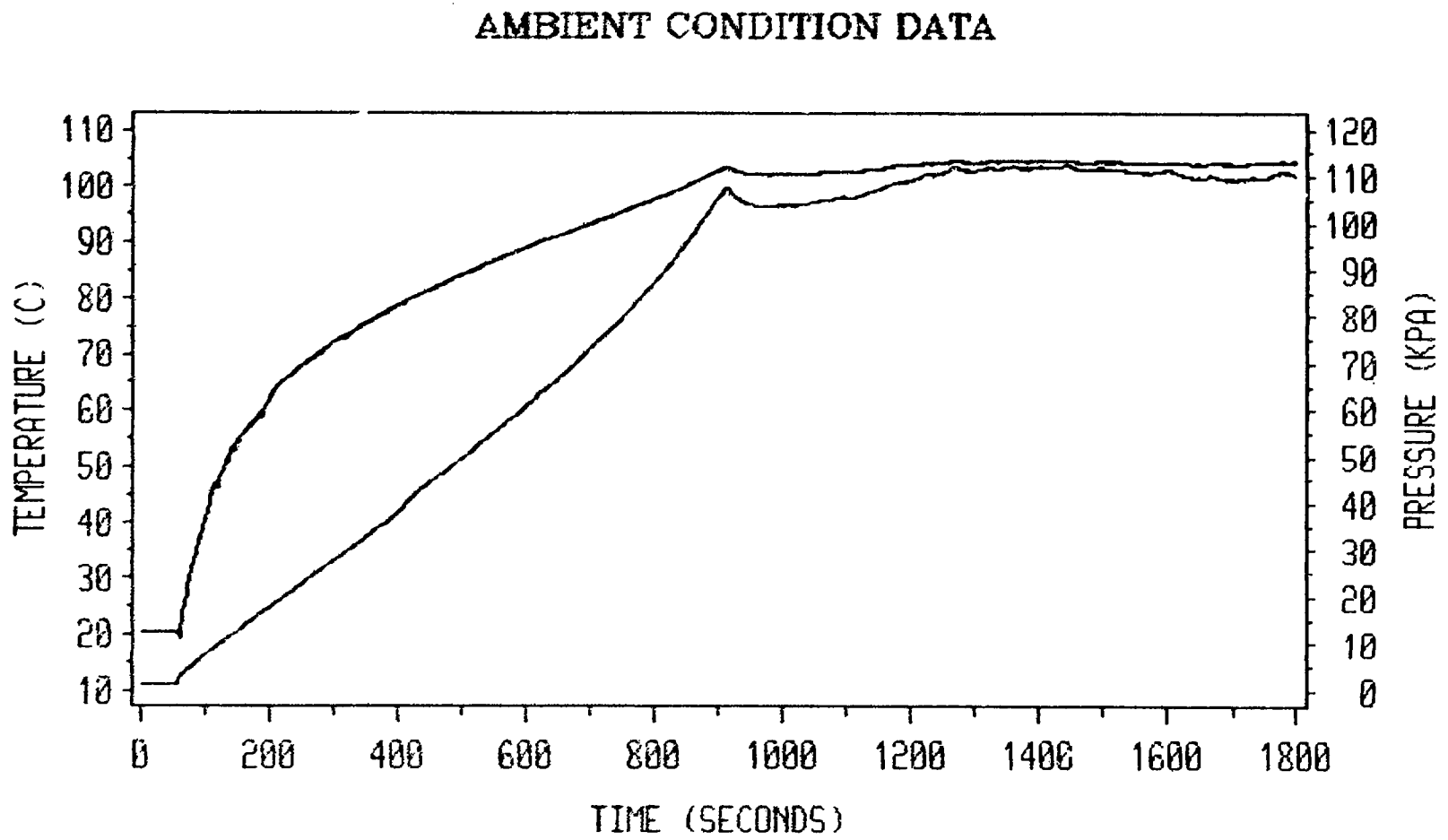
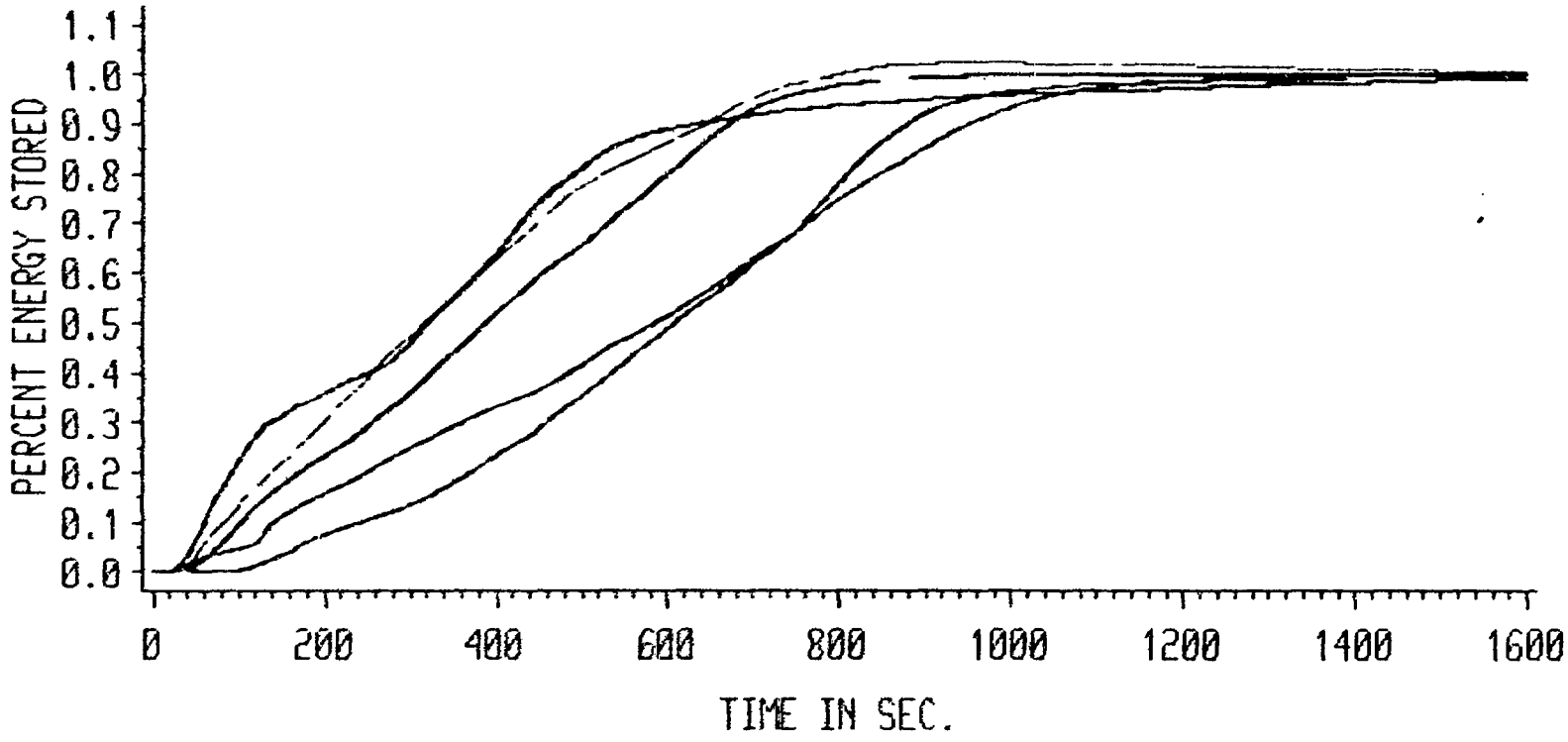


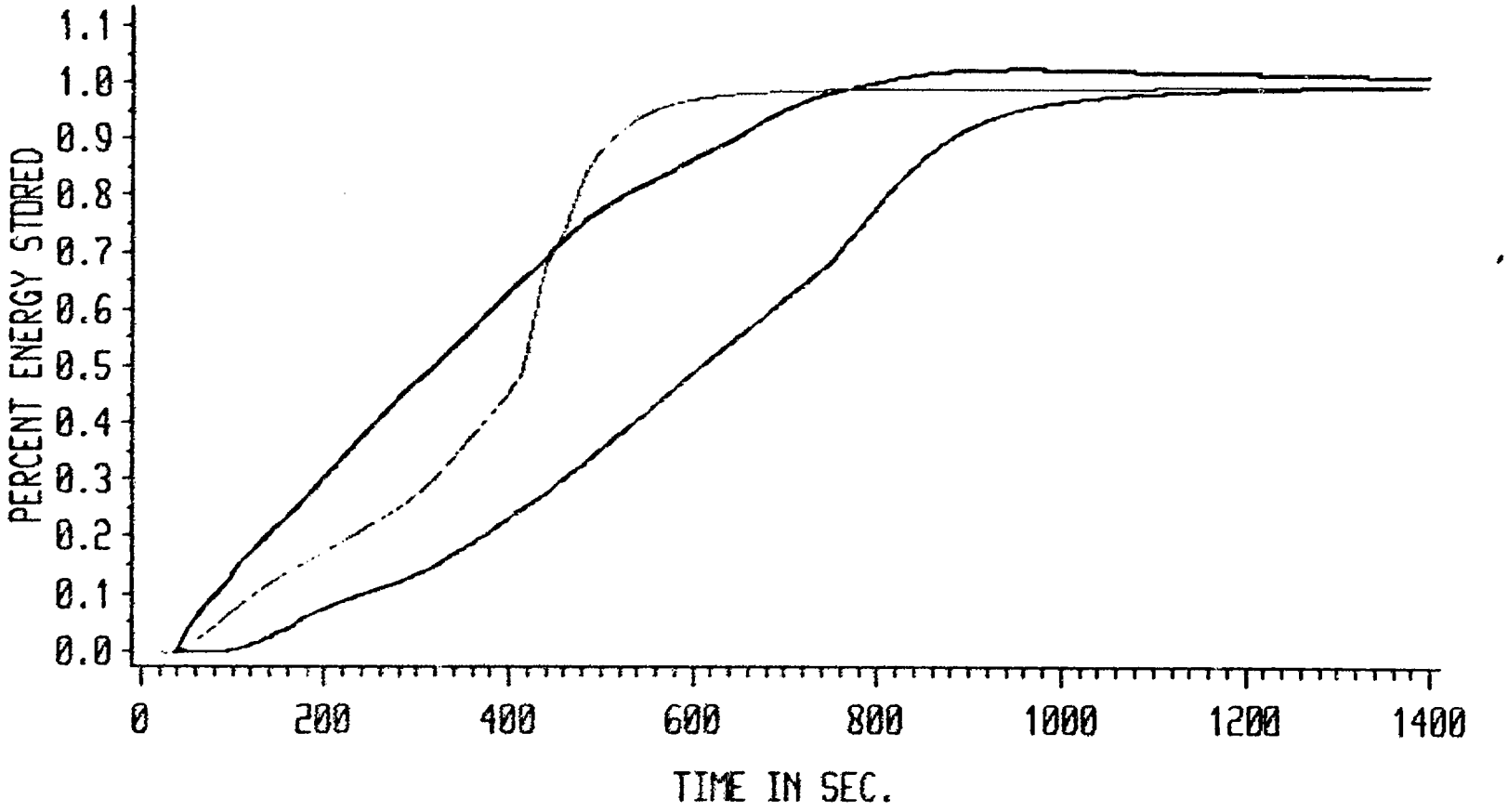
FIG. 27 CMS FN-52. PLOTS OF TEMPERATURE AND PRESSURE IN SPHERE AMBIENCE AS A FUNCTION OF TIME FOR 3.0" DIAMETER ARRAYED SPHERE DATA AT 1 SEC. MC5600 FN-MAIN302.



3.8" DIAM. ISO. SPHERE--95% IN 853 SEC.S
3.8" DIAM. ARR. SPHERE--1007 SEC.S
2.5" DIAM. ISO. SPHERE --- SEC.S
2.5" DIAM. ARR. SPHERE--688 SEC.S

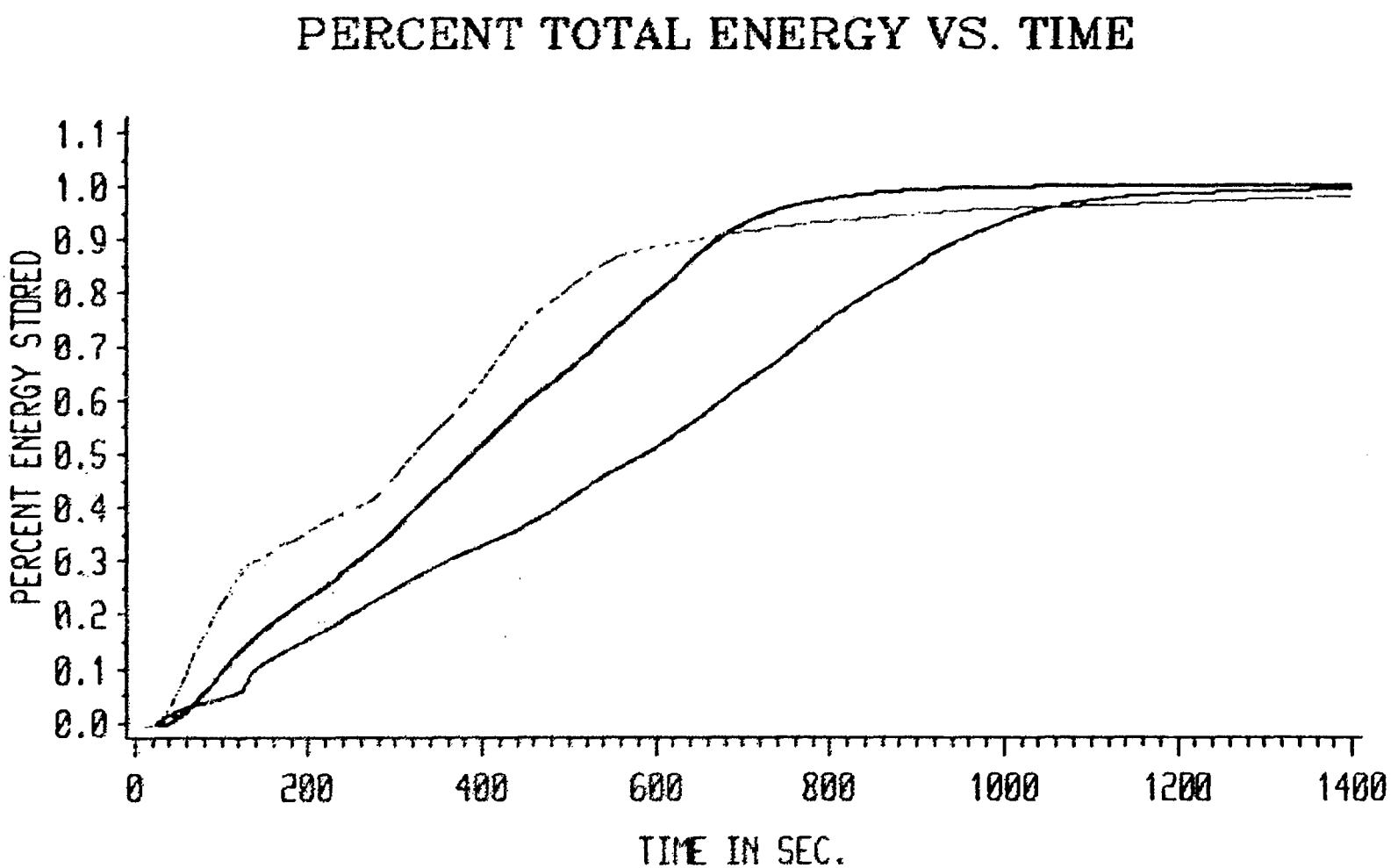
2.8" DIAM. ARR. SPHERE--866 SEC.S

FIG. 28 PLOTS OF THE PERCENTAGES OF THERMAL ENERGY STORED FOR EACH DATA SET VERSES TIME. CMS FN-94.



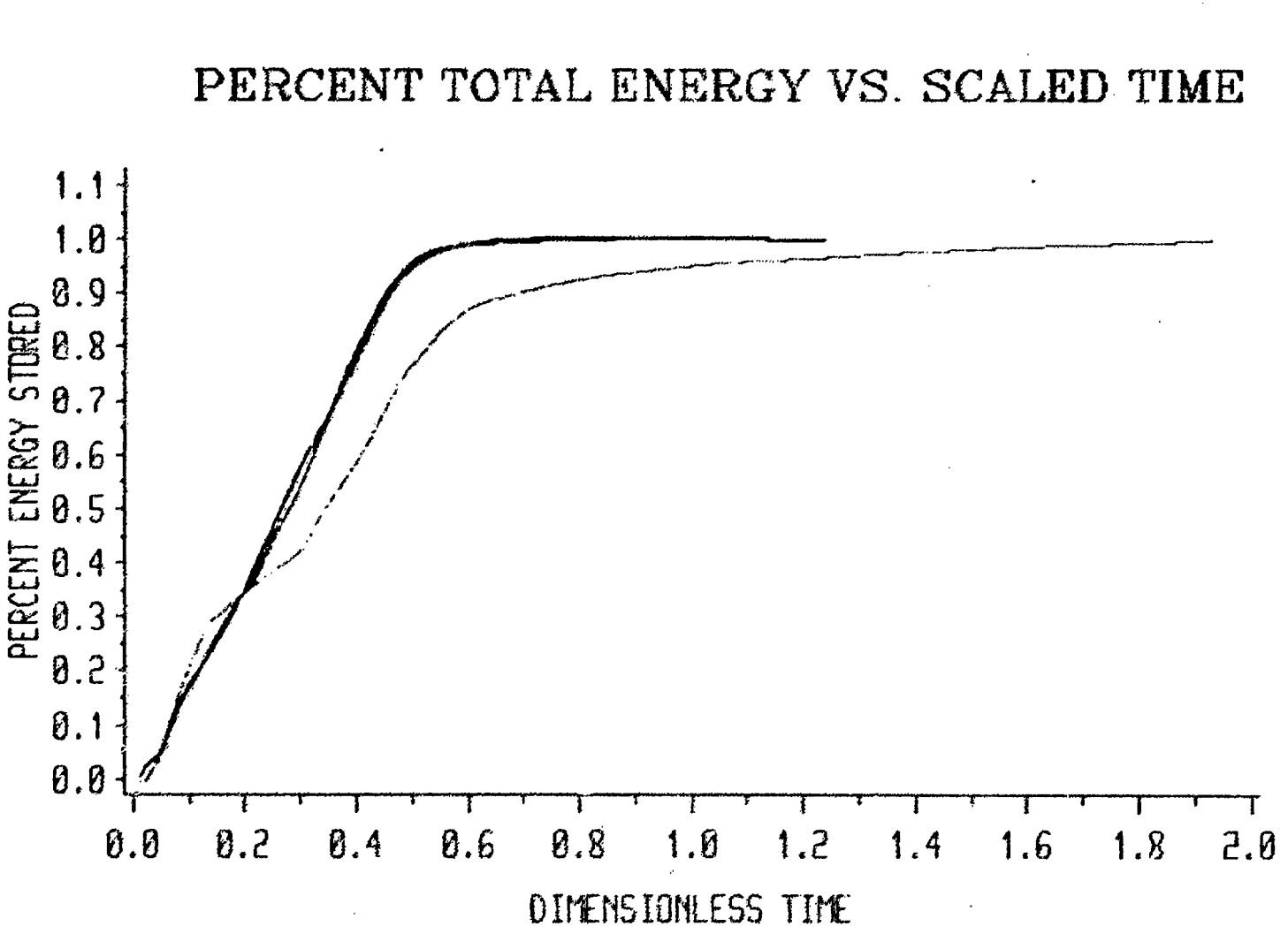
3.8" DIAM. ISO. SPHERE--95% IN 853 SEC.S
2.5" DIAM. ISO. SPHERE--659 SEC.S
2.8" DIAM. ISO. SPHERE--516 SEC.S

FIG.29 PLOTS OF THE PERCENTAGES OF THERMAL ENERGY STORED FOR EACH DATA SET VERSES TIME.
CMS FN-94.



3.8' DIAM. ARR. SPHERE--95% IN 1807 SEC.S
2.5' DIAM. ARR. SPHERE--682 SEC.S
2.8' DIAM. ARR. SPHERE--[illegible] SEC.S

FIG.3D PLOTS OF THE PERCENTAGES OF THERMAL ENERGY STORED FOR EACH DATA SET VERSES TIME.
CMS FN-94.

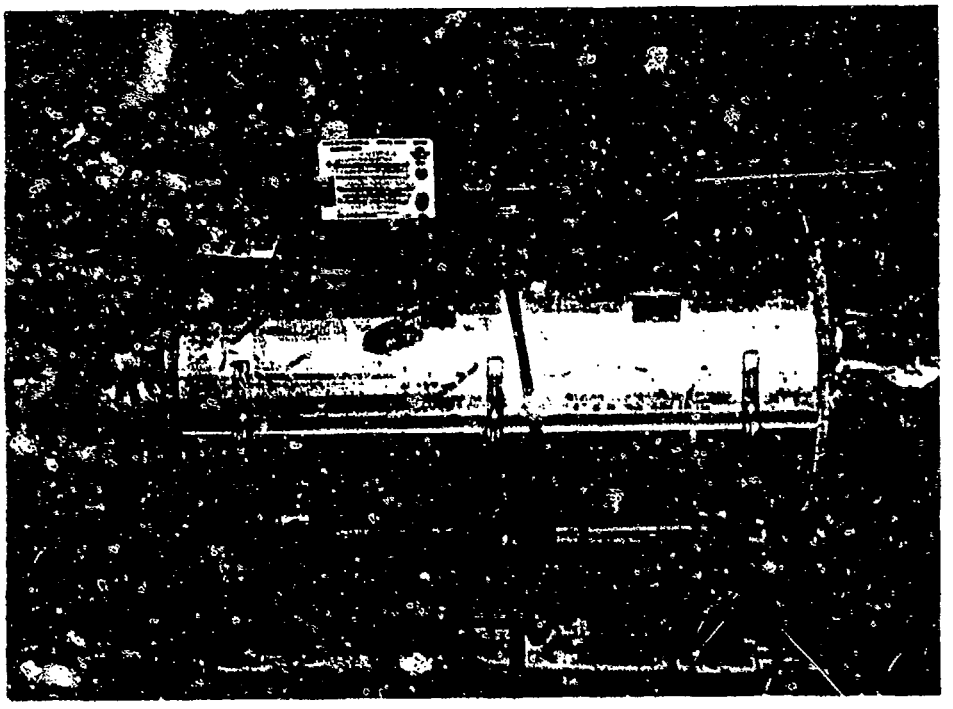


3.8' DIAM. ARR. SPHERE--95% IN 1807 SEC.S
2.5' DIAM. ARR. SPHERE--688 SEC.S
[illegible]

FIG.31 PLOTS OF THE PERCENTAGES OF THERMAL ENERGY STORED FOR EACH DATA SET VERSES TIME. CMS FN-94.



3.1 System Pump



3.2 Preheater



3.2 super heater



3.3 Testing Chambers



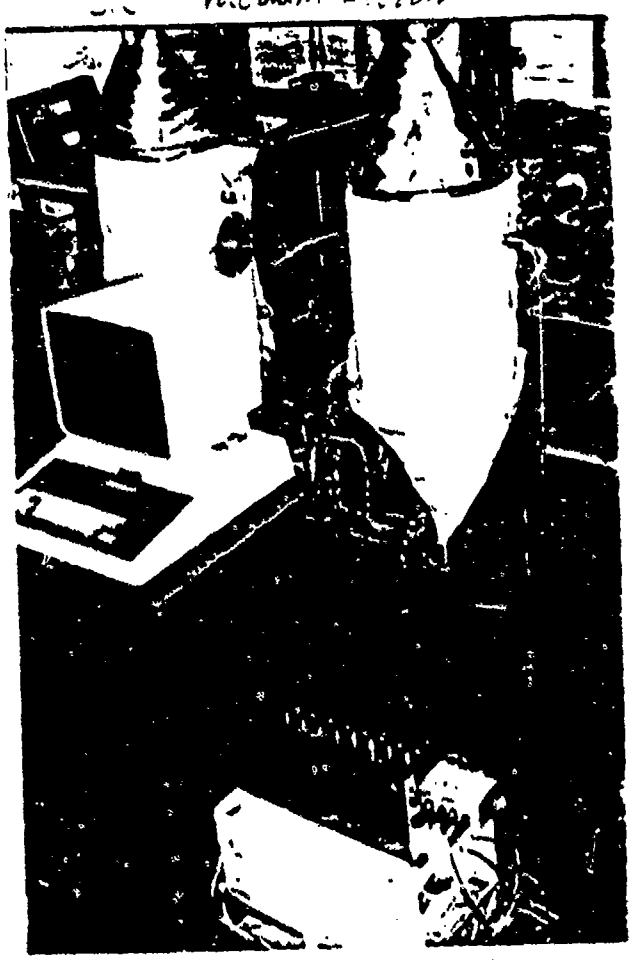
3.3 Aluminum Insert in Keel Chamber



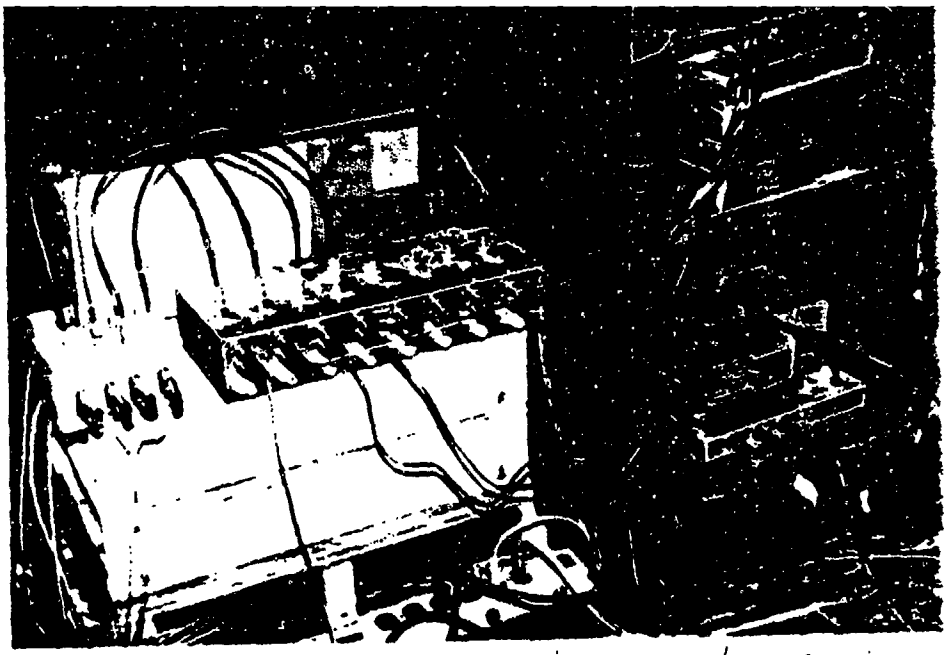
3.4 Condenser



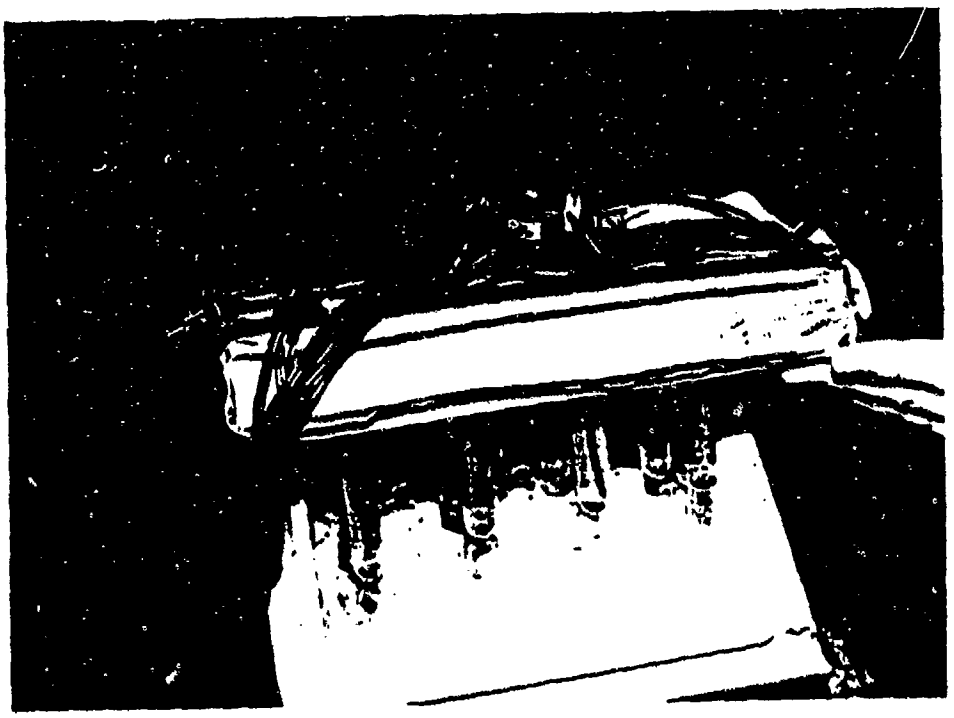
3.5 Vacuum System



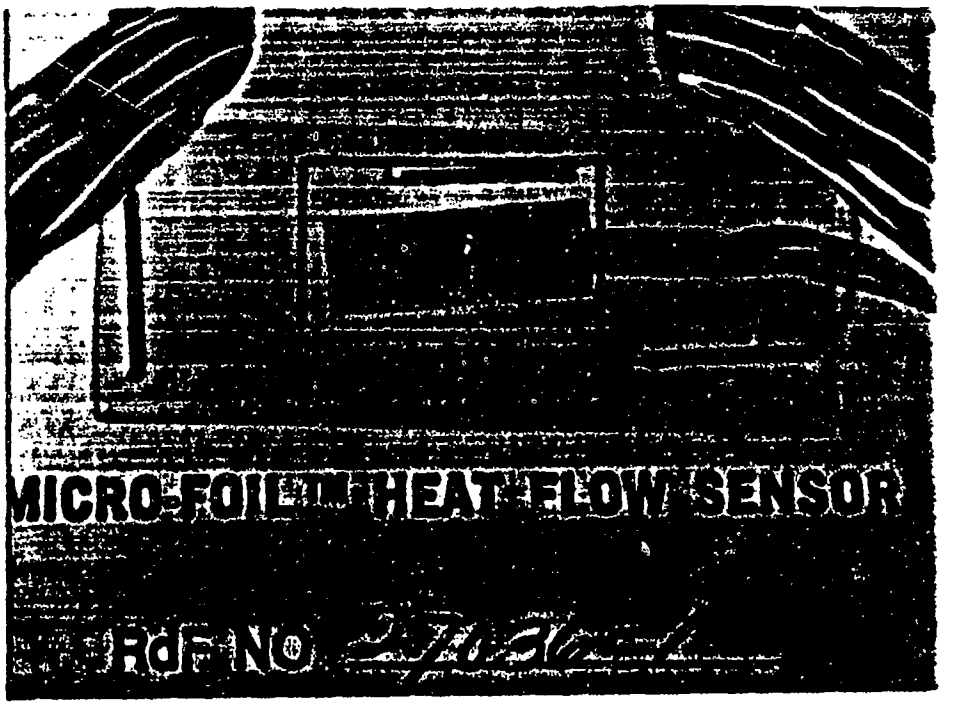
3.6 [illegible]



3.6 Amplifier / Filter / D.C. voltage Supply and Ice bath



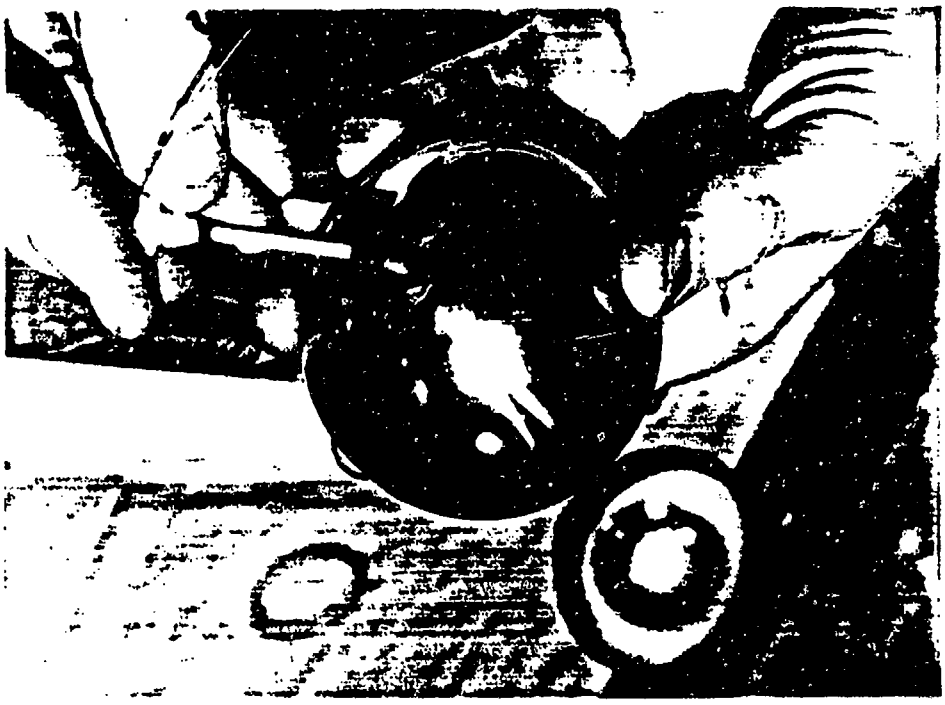
3.8 Ice Bath



3.9 Heat Flux Transducer



3.9 Heat Flux Transducer and Encapsulated Spheres



3.10 Center Temperature Probe



3.10 center Temperature probe

Appendices can be obtained from

Universal Energy Systems, Inc.

1986 USAF-UES Minigrant Program

Sponsored by the

Air Force Office of Scientific Services

Administered by

Universal Energy Systems, Inc.

Final Report

Long-term Potentiation in Interneurons in the Mammalian Brain

Prepared by:	Brenda J. Claiborne
Academic Rank:	Assistant Professor
Department and	Division of Life Sciences
University:	University of Texas, San Antonio
Research Location:	University of Texas, San Antonio
Date:	January 25, 1988

Long-term Potentiation in Interneurons in the Mammalian Brain

by

Brenda J. Claiborne

ABSTRACT

Interneurons in the hilar region of the rat dentate gyrus were examined using *in vivo* and *in vitro* techniques. First, hilar interneurons on one side of the brain were labeled by making *in vivo* injections of fluorescent latex microspheres into the contralateral dentate gyrus. Second, pyramidal and granule neurons in the slice preparation were labeled intracellularly with the dye Lucifer Yellow. Results showed that the axons of these neurons were labeled over greater distances than is possible with the intracellular marker horseradish peroxidase. And third, although healthy granule and pyramidal neurons were encountered in the slices from rats, hilar interneurons did not appear to be viable, even when various changes were made in the standard slice maintenance procedures. As discussed, this finding has important implications for all studies of hippocampal circuitry that make use of the rat slice preparation.

Acknowledgements

I would like to thank the Air Force Office of Scientific Services for funding the research discussed here. I would also like to Dr. David Terrian, USAF School of Aerospace Medicine, Brooks Air Force Base, Texas, for the opportunity to work with him as a Summer Faculty Research Fellow during the summer of 1987. The results that we obtained during that summer led directly to the work reported here. I would also like to thank the students working in my laboratory on this project over the past year. Laura Rihn prepared and maintained the slices and did some of the intracellular injections of Lucifer Yellow. Sue Levey and Rob Connaughton worked on the *in vivo* latex microsphere injections, and Marlana Lehman took part in the photomicrography. They are all to be commended for their excellent work. I would also like to thank the staff at Universal Energy Systems for their prompt, courteous and efficient service in administering the Minigrant Program.

I. Introduction

One of the goals of neurobiological research is to understand the anatomical, physiological and biochemical basis of learning and memory in the mammalian brain. A wealth of data now implicates one cortical region, the hippocampal formation, in short-term memory processing. For example, bilateral damage to the human hippocampal formation produces generalized anterograde amnesia (Milner, 1970; Penfield and Mathieson, 1974; Woods et al., 1978), lesions limited to the hippocampus in monkeys cause memory impairment (Zola-Morgan and Squire, 1986), and dorsal hippocampal removal in rodents results in learning deficiencies (Olton et al., 1979; Kesner, 1985), as does hypoxic ischemic damage restricted to the hippocampus (Volpe and Hirst, 1983; Volpe et al., 1984).

Although it appears to play an important role in the complicated processing of memory, the mammalian hippocampal formation has a relatively simple structure. It consists of three principal neuron types, the granule cells of the dentate gyrus, the pyramidal neurons of the CA3 region of the hippocampus proper, and the pyramidal cells of the CA1 region, and, in addition, contains several types of interneurons (Swanson, 1983).

The synaptic connections between the principal neurons have been well described and of particular interest is the synaptic response known as long-term potentiation (or LTP). LTP is thought to be the cellular basis of memory: it is

characterized by a very long-lasting (up to months) increase in the magnitude of a postsynaptic response following brief periods of intense stimulation of the presynaptic pathway (Teyler and DiScenna, 1984). It has been observed in the CA3 pyramidal cells in response to intense stimulation of the dentate granule cells and in the CA1 neurons in response to intense stimulation of the CA3 neurons (Swanson et al., 1982). The subcellular mechanisms underlying LTP are unknown, but both presynaptic and postsynaptic models have been proposed (Lynch and Schubert, 1980; Bliss and Dolphin, 1982; Swanson et al., 1982; Barrioneuvo and Brown, 1983; Lynch and Baudry, 1984).

Less well understood, however, are the synaptic responses of the interneurons in the hippocampal formation. Some are thought to be involved in inhibitory feedback loops, whereas others are thought to make excitatory feedback connections onto the principal neurons (Swanson, 1983). One region of the hippocampal formation, the dentate hilus, contains a heterogeneous group of interneurons that receive inputs not only from the dentate granule neurons, but also from various modulatory neurons in other brain regions. These hilar neurons have been described anatomically (Amaral, 1978), but only two recent reports exist on their physiological properties (McNaughton et al., 1983; Scharfman and Schwartzkroin, 1987). Their connections are of interest because anatomical studies have shown that most project "back" to the granule neurons; the

granule neurons serve as the first step in information processing in the hippocampal formation and hence influence signal flow through the hippocampal formation.

The hilar neurons, therefore, have the potential for modulating activity in the entire hippocampus, and it would be of particular interest to know if they exhibit LTP. If so, the hilar interneurons could modulate information processing over long periods of time and thereby play a pivotal role in the formation of memory. Until recently it has not been technically possible to reliably record the responses of the hilar cells simply because it has not been possible to make repeated intracellular recordings from any group of interneurons in the mammalian cortex. Several years ago, however, Katz and his colleagues (Katz et al., 1984) devised an in vivo method of labeling the cell bodies of selected cortical neurons with a unique fluorescent marker. After the neurons are labeled, slices of the brain region containing the labeled cells are made, individual neurons are visualized and intracellular recordings can be made. The fluorescent marker used by Katz is unique because it is non-toxic in the presence of ultraviolet light: toxicity caused by illumination has prevented the use of other fluorescent dyes for dual labeling and recording studies.

II. Objectives of the Research Effort

The overall objective of the research effort was to use the method pioneered by Katz to determine if

interneurons in the rat dentate hilar region exhibit LTP following stimulation of the granule neurons. The project was divided into four studies, each with a specific goal.

The specific goals were:

A. To establish the correct stereotaxic coordinates for injecting latex microspheres into the hilar region of the rat dentate gyrus.

B. To determine if unlabeled hilar interneurons are viable in the rat in vitro slice preparation.

C. To establish the parameters for filling neurons in the slice preparation with Lucifer Yellow.

D. To make intracellular recordings in the slice from hilar interneurons labeled with latex microspheres.

III. Experiments Performed

A. Microsphere Injections In Vivo

Although it is possible to record from a small sample of interneurons by making random penetrations into cortical tissue, reliable recordings can only be made when it is possible to visualize the neuronal cell bodies. Therefore, the first goal was to label the hilar interneurons using microspheres. Because most hilar neurons project contralaterally (in addition to projecting back to the granule cells), they can be labeled by injecting microspheres into the contralateral dentate gyrus in vivo. Preliminary studies had shown that the axon terminals of the hilar interneurons take up the microspheres from the

surrounding tissue and transport them to their respective neuronal cell bodies in the contralateral hippocampus. The cell bodies of the interneurons are then visible in the slice preparation under ultraviolet epi-illumination. This part of the study was designed to establish the correct stereotaxic coordinates for making small injections of latex microspheres into the rat dentate gyrus in vivo.

Animals: Sprague-Dawley rats, either male or female, from Harlan Sprague-Dawley, Inc. were used for this part of the study. Males and females were between the ages of 4 and 8 weeks and weighed between 154 and 401 grams. Animals used in all portions of this project were housed in the Animal Care Facility at the University of Texas at San Antonio and all procedures pertaining to their care and use met the standards of the UTSA Animal Care and Use Committee and the National Institutes of Health Guide for the Care and Use of Laboratory Animals.

Anesthesia: Animals were deeply anesthetized with Nembutal injected intraperitoneally. A standard dosage of 0.05 mg/g body weight (American Association for Laboratory Animal Science) was used at the beginning of the study. This dosage was not sufficient to produce deep anesthesia, however, and it was usually necessary to give additional injections of one tenth of the original dose at 10 minute intervals before beginning surgery. In later experiments,

the initial dosage was increased to a level just below that which produced signs of bronchial spasm in the animals, but which eliminated the need for further injecitons.

Fast Green: Because the fluorescent latex microspheres were expensive, the dye Fast Green was used to establish the proper injection coordinates. Fast Green was injected through glass capillary pipettes (outer diameter of 1.2 mm, inner diameter of 0.6 mm) made on a vertical pipette puller. It was found that a saturated solution of Fast Green in 2 M NaCl, filtered before use, caused the least electrode blockage. When the pipettes had been filled with the Fast Green solution, the tips were broken back to final resistances of between 1 and 2.5ΩM . Hyperpolarizing current (20 micro-amps for 10 minutes) was used to eject the dye from the pipette.

Injection Coordinates: Before surgery, the animals were deeply anesthetized with Nembutal. Their heads were shaved and they were positioned in the sterotaxic apparatus with the proper adjustment of the incisor bar. A midline incision of about 2 cm was made to expose Bregma. The glial layers were resected and a mark made on the skull at the pre-determined coordinates. A burr hole was drilled in the skull at this position, exposing the dura. The coordinates were checked before the injection was done, an incision was made in the dura and the pipette lowered into the brain.

Initially, coordinates from Paxinos and Watson (1986) were used. They were found to be indaequate and were systematically varied. The anterior/posterior coordinate was varied from -6.3 mm (measured from Bregma) to -4.3 mm. The lateral coordinate was adjusted along with the anterior/posterior coordinate and varied from 2.0 to 4.6 mm. A microdrive was used to lower the pipette to the chosen doral/ventral coordinate. Values ranged from 3.2 mm (measured from the dura) to 6.0 mm. To minimize tissue compression, the microdrive was advanced slowly in a series of 1 mm steps interspersed with 5 min rest periods to the desired depth. After the injection, the pipette was allowed to remain in the tissue for 5 minutes before removal.

Injection of Latex Microspheres: When the proper coordinates for making an injection into the hilar region were established, latex microspheres were injected using these coordinates. The microspheres were obtained from Tracer Technologies, were supplied in a concentrated solution in distilled water and were not diluted before use. They were injected with pressure as parameters for iontophoretic injection are not known. Rats were anesthetized and prepared for surgery as described above. A 1 ul Hamilton syringe with a 30 gauge needle was used and 0.1 ul of microspheres were injected.

Histology: To assess the success of the injection, animals were sacrificed immediately after the injection was

completed, when they were still deeply anesthetized with Nembutal. Each animal was perfused transcardially with 0.9% saline, followed by 10% formaldehyde in 0.1 M phsophate buffer. The rat was decapitated, and the brain removed and placed in a solution of 10% formaldehyde with 25% sucrose at 4°C for 24 hours. The brain was blocked and sections between 30 and 60 um cut using a freezing microtome. Sections were mounted on gel-coated slides and, for the Fast Green studies, were stained with Neutral Red. Following the latex microsphere injections, sections were not stained (the microspheres dissolve after long exposures to alcohol or zylene), but were observed with a Nikon microscope using bright field optics, phase optics or ultraviolet epi-illumination with a Nikon G-2A filter.

Results: The following doses of Nembutal, injected intraperitoneally, were found to produce the deepest anesthesia without signs of bronchial spasm. For rats between 150 and 200 grams, 0.055 mg/g, for animals between 200 and 250 grams, 0.06 mg/g and for those between 250 and 400 grams, 0.065 mg/g.

As is shown in Figures 1 and 2, a set of coordinates was obtained that could be used to reliably inject fluorescent microspheres into the dentate hilar regions of rats between 150 and 200 grams. These coordinates were -4.3 mm from Bregma in the anterior/posterior dimension, 2.0 mm laterally and 3.5 mm from the dura in the dorsal/ventral dimension. Correct coordinates were also obtained for rats

weighing between 200 and 250 grams (data not shown). These were -4.8 mm from Bregma, 2.5 mm laterally and 4.0 mm in the dorsal/ventral dimension.

Because the concentrated latex microspheres appear pink under white light conditions, the injection sites were easily seen with brightfield or phase optics (Figure 1 and Figure 2a). They were also clearly seen using ultraviolet, epi-illumination (Figure 2b). The microspheres were confined to the dentate gyrus, and as seen in the micrographs in Figure 1 and 2, sometimes aggregated along the granule cell layer. Under ultraviolet light it was possible to visualize microspheres that were sequestered into individual neurons at the edge of the injection site (Figure 2b). Most importantly, when the contralateral dentate gyrus was examined, labeled neurons were clearly seen in the hilar region.

B. Intracellular Recording and Labeling

The goal of this portion of the study was to determine whether or not the interneurons were viable in vitro under standard recording conditions.

Animals: At the begining of this study, male Sprague-Dawley rats weighing between 130 and 404 grams were obtained from Zivic-Miller. Later, as described below, male and female Sprague Dawley rats weighing between 124 and 326 grams were obtained from another supplier, Harlan Sprague-Dawley, Inc.

Tissue Slice Preparation: Slices were prepared as described previously by others (Skrede and Westgard, 1971; Yamamoto, 1972; Schwartzkroin, 1975; Dunwiddie and Lynch, 1979; Dingledine, 1984) and myself (Claiborne et al., 1986). Rats were deeply anesthetized with ether or Nembutal (25 mg/g), decapitated, and the brains removed and rapidly immersed in ice-cold saline for 15 seconds. The hippocampus was removed and transverse slices from the middle third were cut at a thickness of 400 um with a McIlwain tissue chopper. They were placed in warm, oxygenated saline in a recording chamber consisting of 4 small compartments surrounded by a heated waterbath. The slices were supported on a filter membrane (12 um Nuclepore Corp.) and maintained in a buffered saline solution composed of (in mM): 130 NaCl, 5.0 KCl, 2.0 $MgSO_4$, 1.2 KH_2PO_4, 2.0 $CaCl_2$, 26 $NaHCO_3$ and 10 glucose at pH 7.35. In some experiments, the concentration of $CaCl_2$ was increased to 3 mM (see below). The level of the saline was adjusted such that the surface of the slice was at the saline/air interface. A mixture of 95% O_2 and 5% CO_2 was bubbled through a warm water bath and directed over the slices. For most experiments, the slices were maintained at 32°C for 1 to 2 hours before recordings were started; at that time the temperature was raised to 35°C.

Intracellular Recording: Electrodes with very short shanks were made using a Brown-Flaming pipette puller and filled with filtered 3 M KCl (resistances between 30 and 40

ΩM) or 2.5% Lucifer Yellow (see below). Signals were amplified using a Model IR183 pre-amplifier from Neuro Data Instruments Corporation and displayed on a Tektronix 5113 oscilloscope. To assess the health of each slice, recordings were made in the cell body layer of the hippocampal CA1 region, and in the granule cell layer of the dentate gyrus. The slice was judged to be healthy if more than five neurons with resting potentials of at least -50 mV and action potentials of 60 mV or better were encountered in one vertical traverse through the tissue.

If a slice was healthy, a series of systematic electrode penetrations were made throughout the hilar region in an attempt to pentrate the soma of a hilar interneuron. It was necessary to make numerous penetrations because the interneurons are scattered throughout the hilus and cannot be visualized.

Results: To begin these experiments, Sprague-Dawley rats from Zivic-Miller were used. Hippocampal slices were prepared and maintained and only slices in which recordings had been made from granule or pyramidal neurons were used.

Approximately 50 penetrations per slice were made in slices from about 10 different animals. Although small voltage drops (-10 to -20 mV) were seen, no resting potentials (-40 mV or greater) indicative of neurons were encountered. Often drops of -70 to -80 mV were seen and it was presumed that the electrode tip had penetrated a glial cell: no post-synaptic potentials or action potentials were

seen. Several of these presumed glial cells were injected with Lucifer Yellow and the tissue processed (see below). Results showed that indeed these cells were glia.

I postulated that the occasional small drops of -20 mV represented penetrations into interneurons that were simply not healthy enough in the slice to support an electrode. Hence four parameters were changed: 1) electrodes with higher resistances (smaller tips) were used; 2) the calcium concentration was increased to 3 mM; 3) the temperature was increased to 37°C or, in separate experiments, decreased to 32°C, for the entire recording session; and 4) both male and female rats were obtained from another supplier, Harlan Sprague-Dawley, Inc.

Results showed that neither raising electorde resistances nor increasing calcium concentrations improved interneuron recordings, even though both changes resulted in better recordings from granule and pyramidal neurons. Changing the temperature had no effect on interneurons, neither did changing the sex or source of the rats.

These results indicated that in the in vitro rat hippocampal slice preparation, the hilar interneurons are either not healthy enough to support an electrode penetration or do not have resting potentials that are greater than -10 or -20 mV. The latter possibility, however, was ruled out when attempts were made with Lucifer-filled electrodes to inject dye into the "possible neurons" with -20 mV resting potentials -- following tissue

processing, no Lucifer was ever seen inside a neuron in the hilar region.

These results and their important implications are discussed in the final section of this report.

C. Intracellular Injections of Lucifer Yellow

The goal of this section of the project was to elucidate the correct conditions for filling neurons with Lucifer Yellow in the slice preparation so that the interneurons could be filled with Lucifer and their morphological characteristics correlated with the presence or absence of LTP.

Animals: Sprague-Dawley rats obtained from either Zivic-Miller (males weighing between 103 and 404 grams) or Harlan Sprague-Dawley, Inc. (males or females between 124 and 326 grams) were used.

Lucifer Yellow Labeling: Slices were prepared and maintained as described above except that, in the case of the Harlan rats, most experiments were done with a calcium concentration of 3.0 mM which resulted in more stable intracellular recordings. Electrodes with short shanks were made on a Brown-Flaming puller and were filled with varying concentrations of Lucifer Yellow in differing concentrations of lithium chloride. High Lucifer concentrations gave high electrode resistances and caused difficulties in cell penetration, whereas high salt concentrations resulted in little cell filling (presumalby because the salt was ejected instead of the dye).

Successful results were seen with 2.5% Lucifer dissolved in 1 M LiCl.

Electrode penetrations were made in the pyramidal or granule cell layers of healthy slices and when a satisfactory resting potential was obtained (-50 mV or better), the electrode was allowed to sit in the cell for 5 minutes, during which time the resting potential usually improved by -10 mV. Lucifer Yellow was injected using hyperpolarizing current pulses of 3 to 4 nA with a duration of 250 ms at a rate of 2/sec. Using these parameters, dye was injected into granule neurons for between 5 and 10 minutes, and into pyramidal neurons for between 10 and 15 minutes.

Tissue Processing: Following the intracellular injection of Lucifer Yellow, the slices were left in the recording chamber for at least 1 hour to allow the dye to diffuse throughout the neurons. Slices were then removed from the chamber and fixed in 10% formalin in 0.1 M phosphate for 10 to 16 hours at 4°C. They were washed in buffer, dehydrated in ethanol and mounted and cleared in methyl salicylate. The filled cells were examined using ultraviolet epi-illumination with a Nikon B-2A filter and photographed.

Results: A solution of 2.5% Lucifer Yellow in 1 M lithium chloride was found to give the best results. Using electrodes filled with this dye solution, both granule and pyramidal neurons were easily penetrated in the slice, and

after the injection procedures described above, were well-filled with Lucifer Yellow.

The morphologies of the filled pyramidal and granule neurons were comparable to those seen following injections of the non-fluorescent marker, horseradish peroxidase (Claiborne et al., 1986). (Micrographs of the Lucifer-filled neurons were not included in this report because they do not reproduce well in a xerox copy.) The dendrites were filled out to their tips and using a Neofluor 20x objective, spines could be seen along all dendritic branches. Axon branches, or collaterals, were easily seen, as were boutons along their lengths. Surprisingly, the axons of both the granule and pyramidal neurons were filled over distances of approximately 1 to 2 mm, whereas the same axons usually only fill for distances of 500 um with horseradish peroxidase (Claiborne et al., 1986).

These results indicate that it is possible to routinely fill neurons in the slice preparation with the dye Lucifer Yellow and that once intracellular recordings of hilar interneurons are obtained, the interneurons can easily be filled for morphological characterization.

D. Recordings from Labeled Hilar Interneurons

The goal of this part of the study was to combine the techniques described above and 1) to visualize a labeled interneuron cell body using ultraviolet, epi-illumination, 2) while observing it penetrate it with a Lucifer-filled electrode, and 3) record its response to intense

stimulation of the granule neurons and then label it by dye injection. Because of the problems encountered in maintaining healthy interneurons in the slice preparation, this part of the study has not been done yet. As discussed below, this portion will be completed if a slice preparation with physiologically competent hilar interneurons can be developed.

IV. Discussion and Recommendations

Results of the studies described above indicate that hilar interneurons can be labeled by injecting latex microspheres into the contralateral dentate gyrus of rats, that neurons in the slice preparation can be routinely filled with the fluorescent dye Lucifer Yellow, and furthermore, that Lucifer fills axons over greater distances than does the marker horseradish peroxidase. It appears, however, that in the rat in vitro slice preparation, hilar interneurons are not viable when maintained under recording conditions.

The latter finding was unexpected, considering that healthy pyramidal and granule neurons are routinely penetrated in hippocampal slices. Recently it came to my attention that other investigators, working independently in Israel, have also found that hilar interneurons are not physiologically healthy in hippocampal slices prepared from rats (L. Katz, personal communication; October, 1987). These investigators labeled the interneurons by making *in vivo* microsphere injections in the contralateral

hippocampus and then, under visual control, attempted to penetrate the interneuron cell bodies with electrodes. They also failed to obtain resting potentials, even though they were able to inject some Lucifer into the neurons. (They argue that the microspheres were not toxic to the neurons because nearby pyramidal neurons were also labeled with the microspheres and these cells exhibited normal resting potentials and action potentials.)

Hence it appears that hilar interneurons are not functional in the in vitro rat slice preparation and therefore that the neuronal circuitry in the slice is incomplete. This has far-reaching implications for all of the studies that have been done on hippocampal circuitry using the rat slice preparation and deserves further investigation.

In contrast to the above findings in the rat hippocampal slice preparation, recent work reported at the Annual Society for Neuroscience Meeting in November, 1987, suggests that the hilar interneurons in slices prepared from guinea pigs are physiologically healthy (Scharfman and Schwartzkroin, 1987). The investigators performing this study used standard procedures for preparing and maintaining the slices, and by making a series of penetrations in the hilus (as described above in this study) were able to record from and fill a number of hilar cells. The cells exhibited normal neuronal resting potentials, action potentials and post-synaptic potentials

in response to stimulation at various sites in the hippocampus and dentate gyrus. There was nothing in the report (or that came up in a discussion between myself and one of the authors) to suggest that the technical aspects of their study were significantly different: the only difference appears to be that guinea pigs were used instead of rats.

Therefore, the second part of this project in which the viability of hilar interneurons in tne slice was determined should be repeated using slices prepared from guinea pigs. If the hilar interneurons are viable in these slices, then, with adequate funding, the last part of the project can proceed. It would also be of interest to try other variations in the rat slice preparation in an attempt to find the conditions necessary to maintain hilar interneurons. For example, longitudinal slices might be cut or serum supplements added to tne recording saline.

References

Amaral, D.G. (1978) J. Comp. Neurol. 182: 851-914.

Barrionuevo, G. and T.H. Brown (1983) Proc. Natl. Acad Sci. USA 80: 7347-7351.

Bliss, T.V.P. and A.C. Dolphin (1982) Trends in Neurosci. 5: 289-290.

Claiborne, B.J., Amaral, D.G. and W. M. Cowan (1986) J. Comp. Neurol. 246: 435-458.

Katz, L.C., Burkhalter, A. and W.D. Dreyer (1984) Nature 310: 498-500.

Kesner, R.P. (1985) In D. S. Olton, E. Gamzu and S. Corkin (Eds.) "Memory Dysfunctions: An Integration of Animal and Human Research from Preclinical and Clinical Perspectives" Ann. New York Acad. Sci. New York: New York Academy of Sciences 444: pp. 122-136.

Lynch, G. and M. Baudry (1984) Science 224: 1057-1063.

Lynch, G. and P. Schubert (1980) Annu. Rev. Neurosci. 3: 1-22.

Milner, B. (1970) In K.H. Pribian and D.E. Broadbent (Eds.) "Biology of Memory" Academic Press, New York and London, pp. 29-50.

McNaughton, B.L., Barnes, C.A. and J. O'Keefe (1983) Exp. Brain Res. 52: 41-49.

Paxinos, G. and C. Watson (1986) "The Rat Brain in Stereotaxic Coordinates", Academic Press, Orlando, Florida.

Penfield, W. and G. Mathieson (1974) Ach. Neurol. 31: 145-154.

Olton, D.S., Becker, J.T. and G.E. Handelmannn (1979) Behav. Brain Sci. 2: 313-365.

Scharfman, H. and P. Schwartzkroin (1987) Soc. for Neurosci. Abstracts 17: 1516.

Swanson, L. (1983) In W. Seifert (Ed.) "Neurobiology of the Hippocampus", Academic Press, New York, pp. 3-19.

Swanson, L., Sawchenko, P.E. and W.M. Cowan (1981) J. Neurosci. 1: 548-559.

Teyler, T.J. and P. DiScenna (1984) Brain Res. Rev. 7: 15-28.

Volpe, B.T. and W. Hirst (1983) Arch Neurol. 20: 436-440.

Volpe, B.T., Pulsinelly, W.A., Tribuna, J. and H.P. Davis (1984) Stoke 15: 558-562.

Woods, B.T., Schoene, W. and L. Kneisley (1982) J. Neurol. Neurosurg. Psychiat. 45: 243-247.

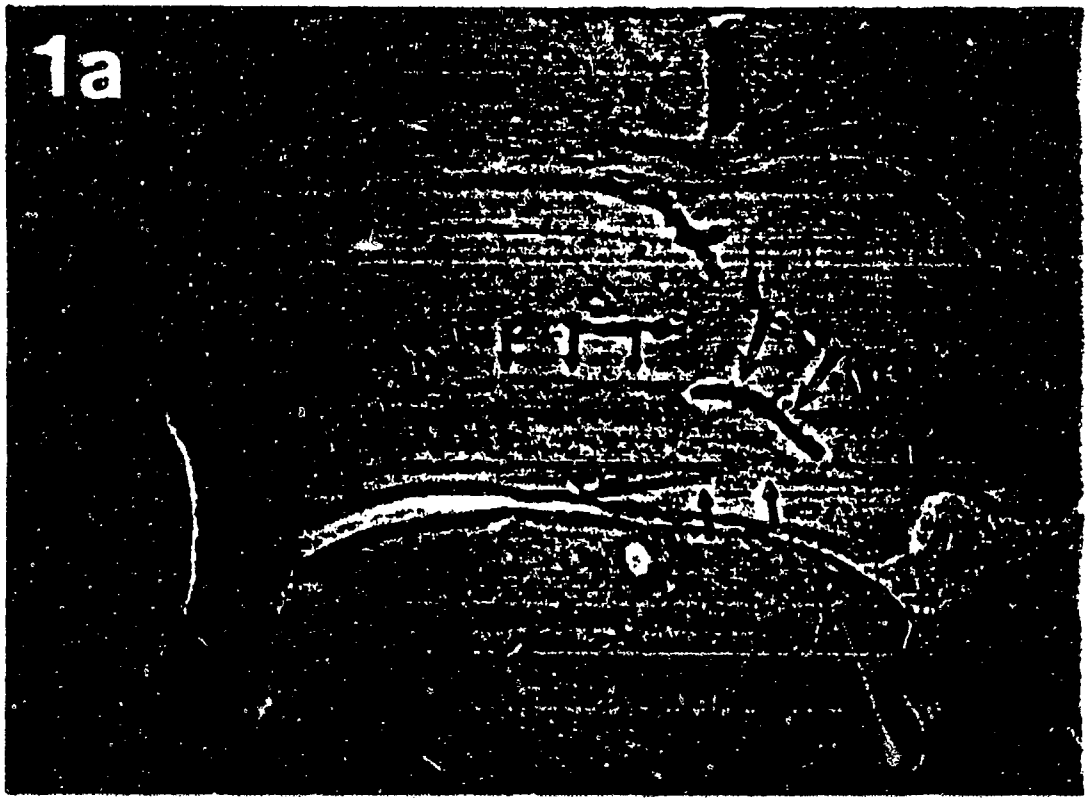
Zola-Morgan, S. and L.R. Squire (1986) Behav. Neurosci. In press.

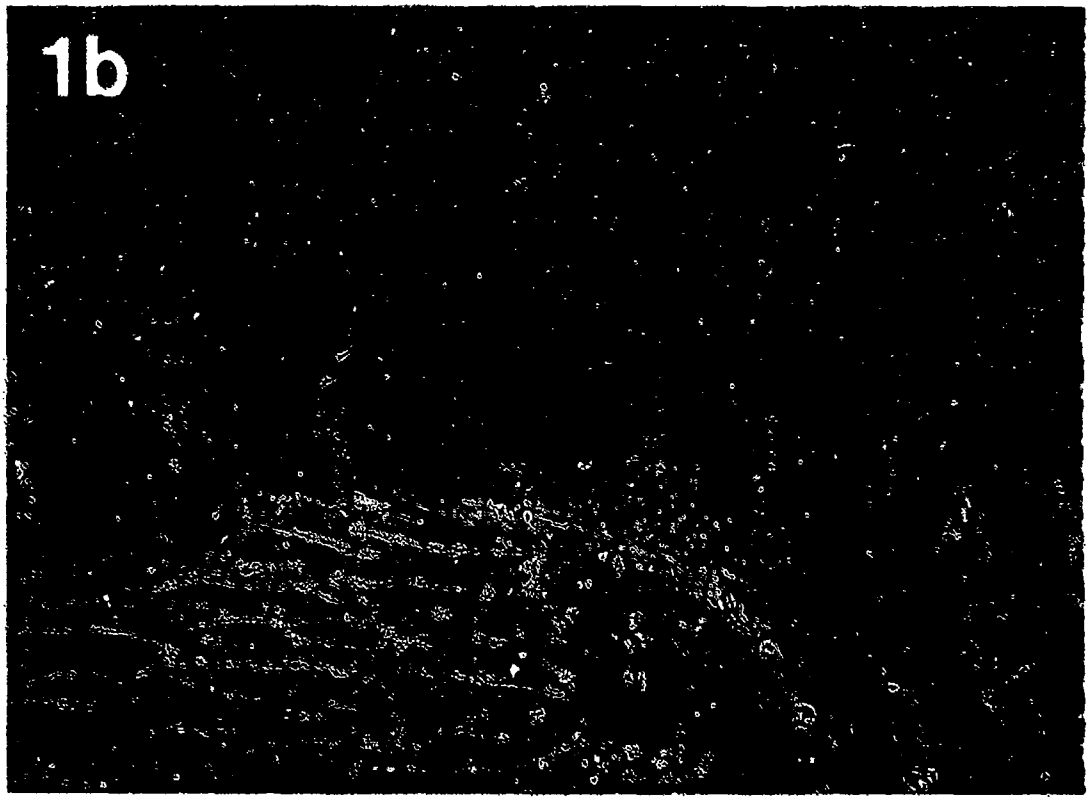
Figure Legends

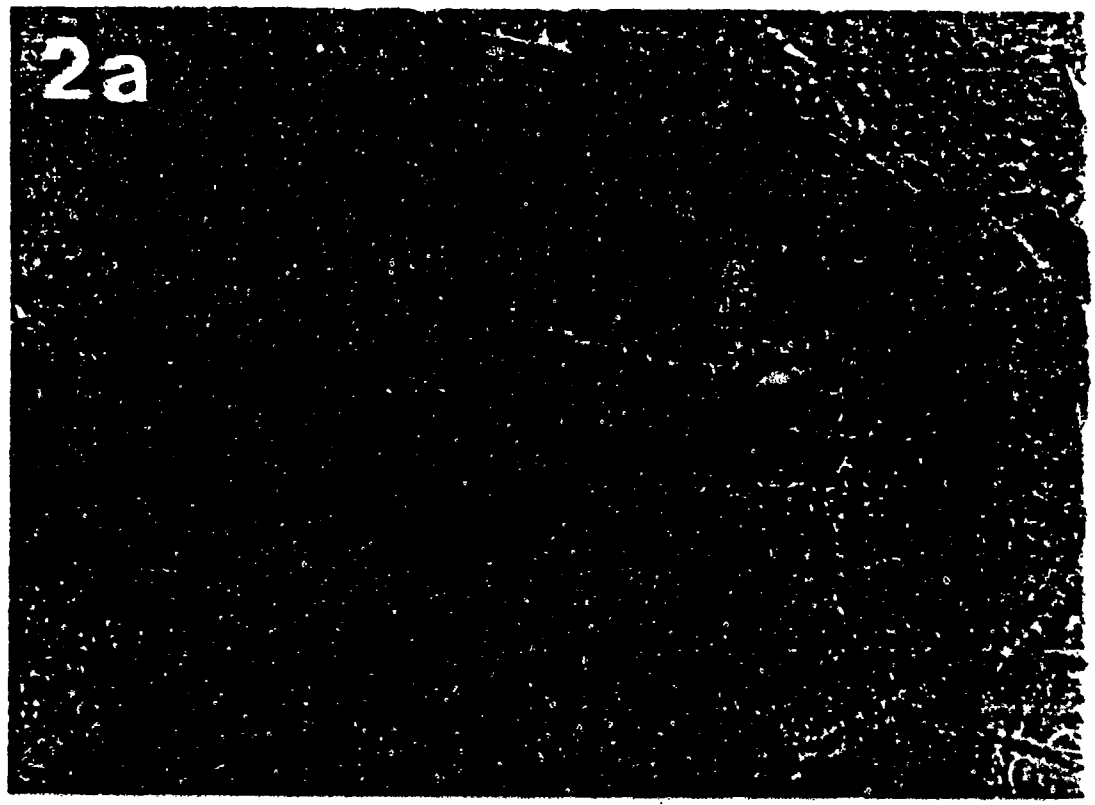
Figure 1: Micrographs of a section from the brain of an animal injected with fluorescent microspheres. These micrographs show microspheres at the injection site in the dentate gyrus as seen with brightfield optics. The injection was made in vivo in an animal weighing less than 200 grams. The low power micrograph on the top (1a) shows the entire hippocampus. The microspheres are aligned along one part of the granule cell layer (large arrows). The smaller arrows delineate the remainder of the granule cell layer. The micrograph on the bottom (1b) shows the injection site at a higher power magnification. The arrows along the bottom outline the granule cell layer surrounding the hilar region (h). Magnifications: a = x23.5; b = x117.5.

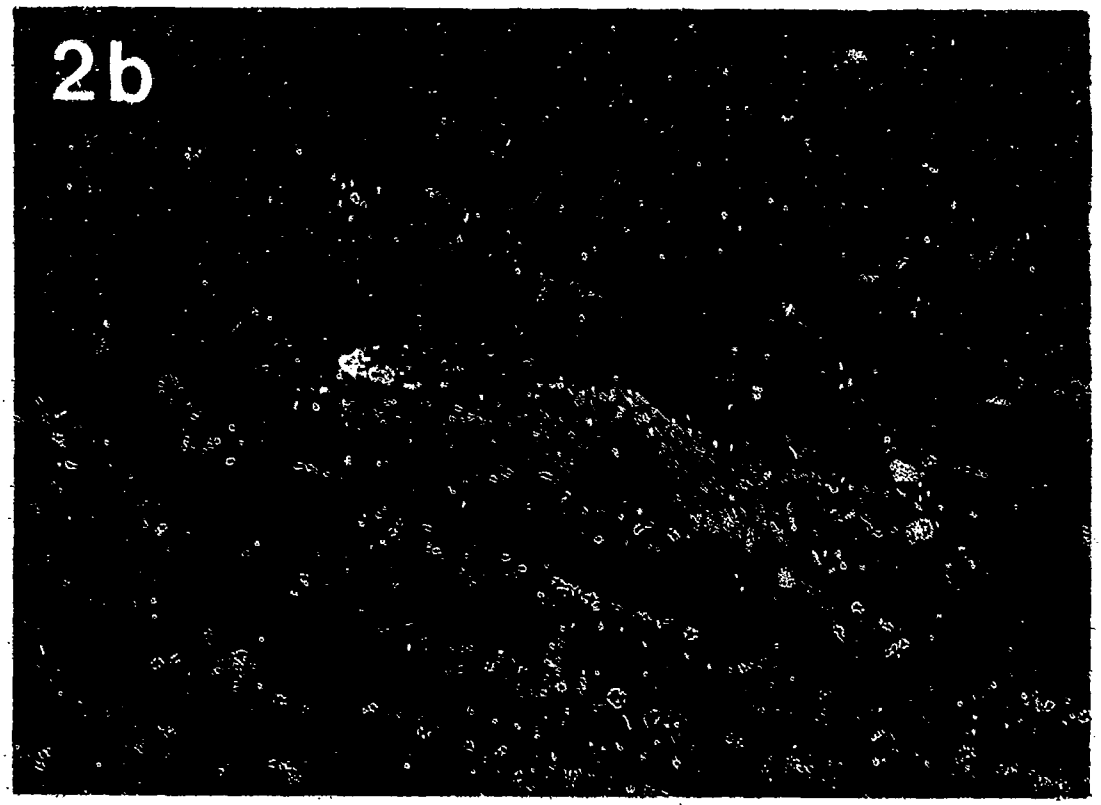
Figure 2: Micrographs of another section from the same animal seen in Figure 1. The top photo (2a) shows microspheres along the granule cell layer as observed under brightfield optics. Arrows outline the granule cell layer around the hilar region (h). The bottom micrograph (2b) was taken of the exact same section as shown in 2a, but ultraviolet epi-illumination was used. It shows the injection site and individual neurons in the hilar region that have sequestered the microspheres (bright spots

underneath the large injection site). Magnification for both a and b = x117.5.


1a


1b


2a


2b

The Development of Image Processing Algorithms for the AFGL Ultraviolet Camera and Other Imaging Systems

Final Report

Submitted to
Universal Energy Systems, Inc.
4401 Dayton-Xenia Rd.
Dayton, Ohio 45432

Contract No.
F49620-85-C-0013
SB5851-0360

P. O. No. S-760-6MG-028

Donald F. Collins
Warren Wilson College
Swannanoa, NC 28778
(704) 298-3325

Abstract

Rocket plume images photographed with the AFGL UV imager on a video system have been analyzed with a video digitizer and an MS-DOS computer. The analysis consists of distortion-free rotations of the plume images to align the plume axis with the horizontal pixel direction, plotting contours and brightness slices of the images, and correcting for atmospheric absorption. Other images of laboratory physics interest have been digitized and analyzed: these include the diffusion of bromine in air, the photography of critical opalescence, Kármán vortices, Bénard cells in clouds, and lunar feature enhancements. In addition the linearity of video cameras was tested, time exposure signal/noise enhancement for low light images and spark discharges were investigated.

Acknowledgements

The author wishes to thank Warren Wilson College for supporting this research by providing released time during the spring and summer of 1987 and by providing space, utilities, and library support for this project. The author also wishes to thank Universal Energy Systems, Inc. and the Air Force Office of Scientific Research for funding the project and the acquisition of the image analysis equipment, Professor N. Rosenberg for making available the image analysis software, Carlos Benavente for photographing the experiments on critical opalescence, and Christos Strergis of the Air Force Geophysics Laboratory for his encouragement and support for this project.

I. Introduction

An image digitizing and analysis system was obtained which digitizes and analyses TV compatible video signals. This consists of a frame digitizer or "frame grabber" with memory (Imaging Technology FG-100/AT), display hardware, and software to perform arithmetic and various image analysis functions on digital images. This equipment provides an economical picture processing system for a single researcher to begin experiments in digital image processing at reasonable cost. A CCD video camera (Fairchild CCD-3002) was also procured for the purpose of photography of virtually any physical system with visible light.

Several applications of this system were developed for complete analysis: rocket plume analysis for the AFGL, molecular diffusion, and critical opalescence. Other potential applications were demonstrated as possibilities for further study: Kármán vortices, Bénard cells in clouds, lunar photography, time-exposure for low-light levels without image intensification, and spark discharge photography.

II. Rocket Plume Analysis

In November of 1986 the Air Force launched a Scout rocket from Vandenberg Air Force Base. This rocket carried the Polar Bear satellite into a successful orbit. The plume was photographed by the AFGL-UV Imager which consists of a UV 3-inch telescope, solar blind photocathode (Cs_2Te on quartz), an MCP intensifier with a P-20 phospho- readout, and a Fairchild video camera (CCD-3002) to generate a video signal for recording and later digitizing. This camera system was calibrated and evaluated by the author during the summer of 1986 as an AFGL Summer Faculty Fellow. The video recording of the Scout plumes was digitized by the author's system (about 1 frame every 5-10 seconds) during the visibility of the rocket plume which extends until 2nd stage burnout. These images were analyzed for size, brightness at the detector, and brightness at the source using the digital image system.

A. Plume Size

A series of images were digitized from a copy of the video tape made by AFGL. Frames were "snapped" manually during playback at normal speed to prevent "sync-tearing" characteristic of VCR's in a still-frame mode. A series of 24 frames were digitized beginning at lift-off and ending

at burn-out of the 2nd stage at 123 seconds. The frames were digitized at full-frame (512 x 480) resolution. A quarter section of the image (240 x 240) which contains the plume was saved on a disk file. One frame which shows the rocket body was zoomed down (alternate pixels) into a 240 x 240 quadrant and saved. This frame was used to calibrate the angular field of view for the camera so that plume sizes could be determined.

Each of the digitized images was rotated so that the plume-axis was aligned horizontally. This rotation had to correct for the 6:5 aspect ratio characteristic of the digitizer and video screens. The rotation algorithm, written by the author, uses the ratio:

$$\frac{k_x}{k_y} = \frac{\text{(horiz. length/pixel)}}{\text{(vert. length/pixel)}} = \frac{6}{5} \quad (1)$$

To calculate the original pixel (x, y) for every pixel (x', y') in the rotated image the rotation matrix is applied:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -(k_y/k_x)\sin\theta \\ (k_x/k_y)\sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} x' \\ y' \end{bmatrix} \quad (2)$$

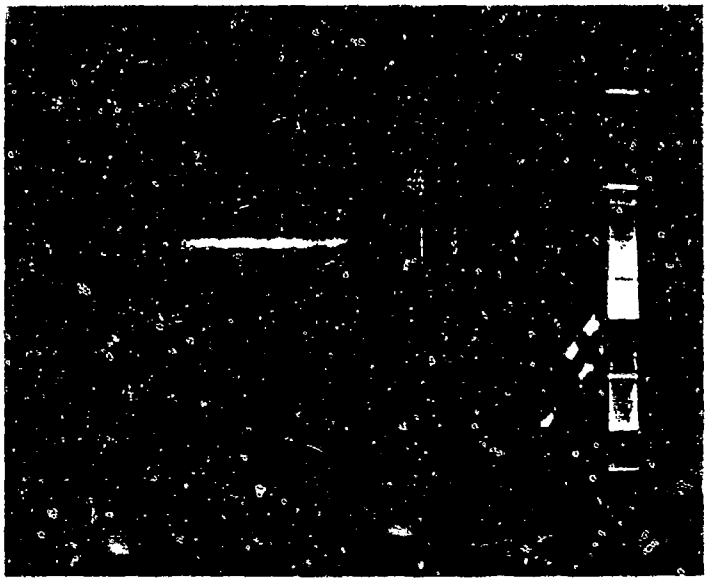
This transformation preserves the circular shape to within 2% for a circular object when rotated. It was checked by

photographing a circular object with the system and rotating 90 deg. The scaling-rotation transformation above was applied in one pass for efficiency and to prevent ragged edges resulting from repeated floating point-integer truncations. Look-up tables were also used for speed.

The rotated images were then filtered with a 9x9 mean, preserving the original pixel value if it exceeded the mean by twice the median of the original neighbor differences. This filter preserves the edges of the image. The images were in general not re-scaled to enhance the contrast. Only the final images of the second stage required contrast enhancement to be seen on the display.

A 64x64 element section was saved for contour plotting by ASYST* software. Most sections saved every third column due to the plume being longer than wide. Photograph 1 shows the filtered false-color image and the section for the contours outlined. Before plotting the contours, the zeroes in the images (from the corners cut-off by rotating) were replaced by the median value of the section image, which is essentially the sky background. Three contours were drawn: 25%, 50%, and 75% of the range between the minimum and maximum of pixel values in the section (Fig. 1). The size (length, diameter,) of the plume (in pixels)

*ASYST software is published by Macmillan Software Co.



Photograph 1. A filtered false-color rotated image of the rocket plume. A 64 x 64 array (every 3rd column and every row) is outlined by the box for contour plotting of Fig. 1.

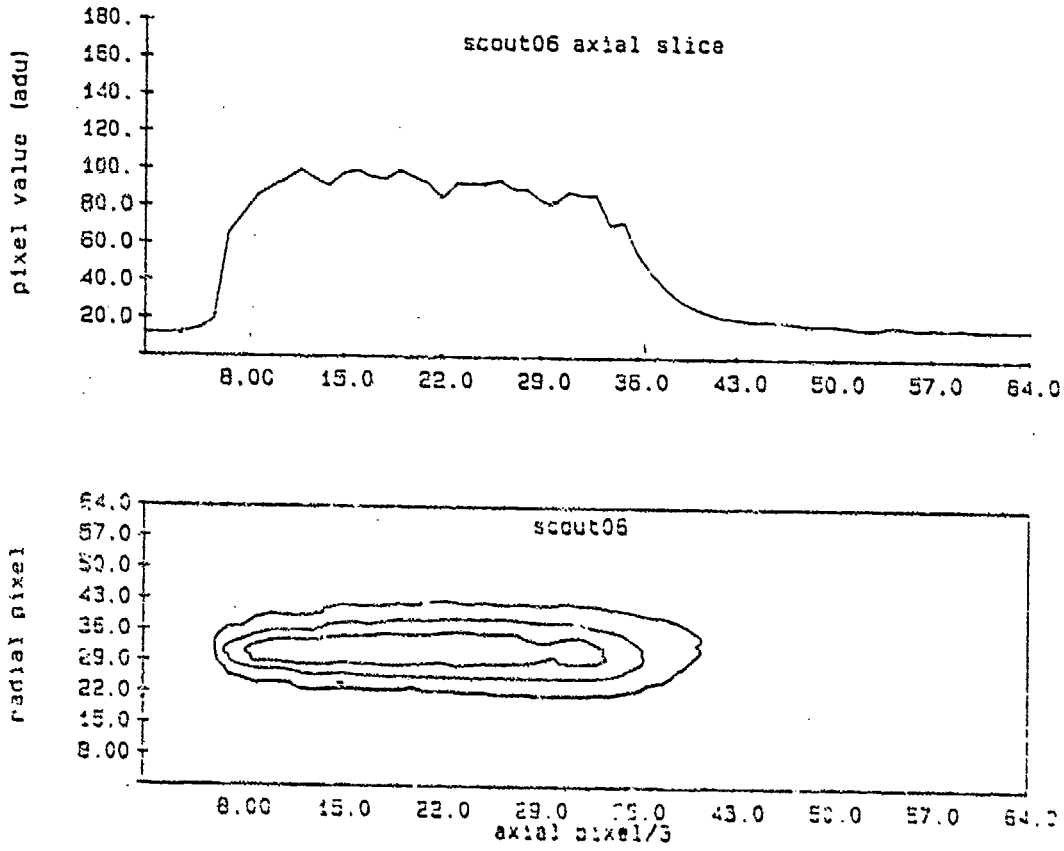


Figure 1. (Bottom): Contour plot of rocket plume showing the 25%, 50%, and 75% brightness contours when the rocket altitude was 27,600 ft. (Top): The pixel values along the axial slice of the image.

was measured from the major and minor axes of the 50% contour. Where the plume image had excessive contrail visible and ran off the picture, the length of the plume was not measured. The diameter was measured through the "head" of the plume - that part where the image was brightest.

An ephemeris of the Scout rocket (made available by the AFGL) gives the geodetic coordinates (lat, long) of the vehicle, the altitude, and the attitude angle of the rocket. These data, the geodetic coordinates of the launch pad and the camera location, enabled the distance from the camera to the rocket to be calculated using geometry. The viewing aspect angle (the angle between the rocket axis and rocket-camera direction) is also calculated. These are displayed in Table I along with the measured plume diameters and lengths. These calculations are needed to calculate the size of the plume from the image size.

The distance from the rocket to camera is obtained by calculating the distance form the (lat, lon) directly beneath the rocket to the (lat, lon) of the camera. The (x, y, z) coordinates of these two points on the surface of a sphere of radius R_{earth} are then calculated and the Pythagorean theorem used to calculate the horizontal distance. The Pythagorean theorem is again used with the

Table I. Partial ephemeris for the Scout rocket with calculated distances to detector and plume sizes.

Time (s)	Altitude (m)	Horiz. dist. (m)	Tot. dist. (m)	View angle (deg)	Plume diam (m)	Plume len (m)
3	160	6205	6222	83.9	3.6	20.8
12	946	6166	6175	91.4	6.4	26.6
15	1393	6136	6186	87.5	5.1	27.7
21	2576	6032	6345	77.7	3.3	25.9
30	5044	5841	7334	59.2	3.9	35.4
37	7558	5834	9072	45.4	3.9	40.6
42	9733	6120	10985	37.3	4.0	46.3
50	14038	7584	15421	27.3	5.2	64.3
58	19644	10923	21946	20.6	9.4	
64	24841	14873	28431	17.1	8.3	
71	31515	20649	37168	14.5	17.6	
78	38246	27058	46353	13.3	18.4	
83	42885	31796	52898	13.0		
88	47595	36900	59742	12.8	23.7	
96	55809	46417	72121	12.1	23.1	
98	58035	49125	75570	11.9	32.3	
102	62789	55023	83029	11.6	27.9	
106	67963	61639	91301	11.3	36.3	
108	70733	65254	95788	11.2	38.1	
110	73636	69081	100523	11.1	46.1	
114	79868	77451	110817	10.9		
118	86709	86826	122277	10.5	44.8	
121	92253	94561	131682	10.2		
123	96353	100329	138674	10.3		
126	101553	107712	147619	10.1		

horizontal distance and the altitude of the rocket above the elevation of the camera as the two legs of the triangle to obtain the total distance from the camera to rocket.

The viewing angle is computed by calculating the direction cosines of both the total distance to the rocket and the orientation of the rocket. The scalar product between these two vectors gives the cosine of the viewing angle. Both the viewing angle and the total distance were checked for qualitative agreement with common sense.

The field of view per pixel on the image analysis system is determined by examining the image at 12 sec. in which the total length of the rocket is visible. This image digitized alternate pixels in order to show the whole rocket in one quadrant of the display area. All other images skip no pixels when sectioning to the 240x240 quadrant. The results of the field of view per digitized pixel, when every pixel is digitized, are: 2.02×10^{-3} deg/pixel horizontal and 1.75×10^{-3} deg/pixel vertical.

The plume lengths and diameters measured from the procedures outlined above using the 50% brightness contour and geometric correction are plotted in Fig. 2 and Fig. 3. Plume lengths were available only for the images when the contrail was absent (altitudes less than 50,000 feet..

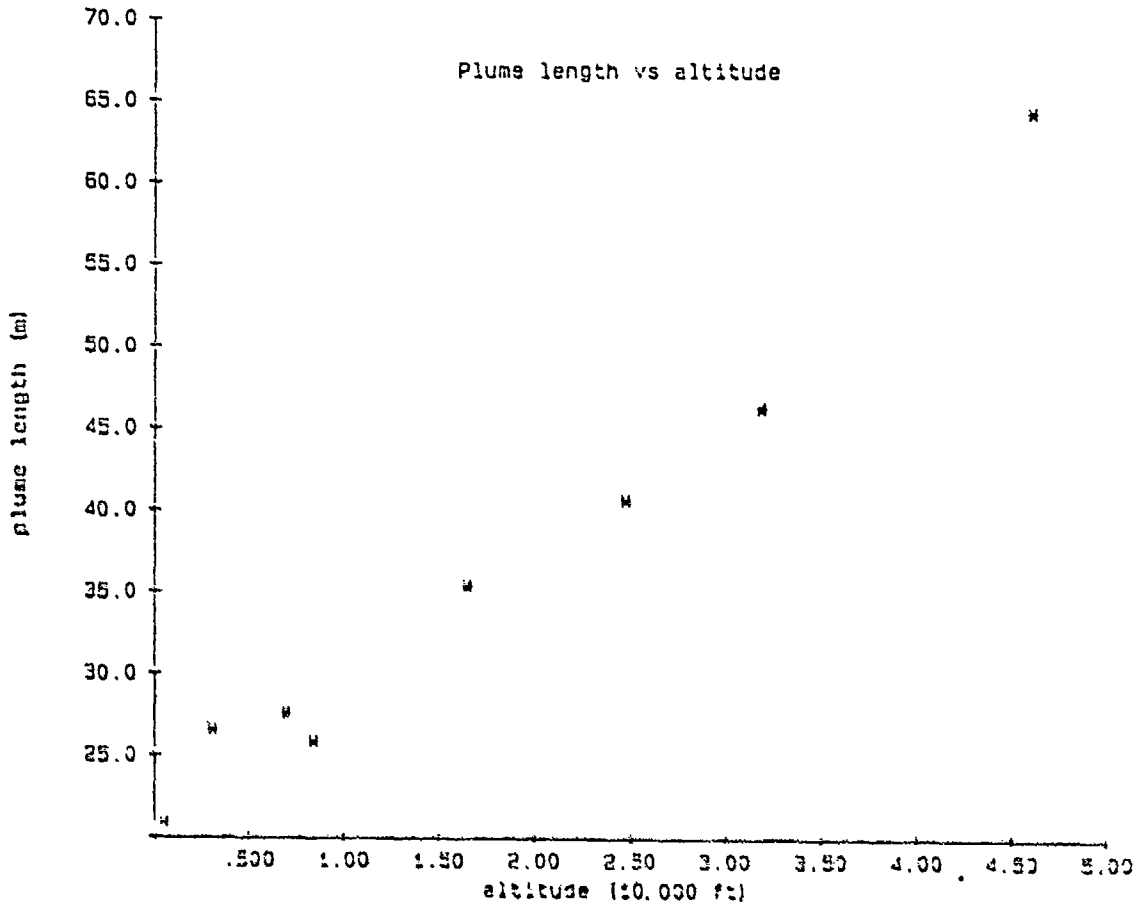


Figure 2. Plume length as measured by major axis of the 50% brightness contour, accounting for the angle between rocket axis and direction to camera, and the angular field of view per pixel.

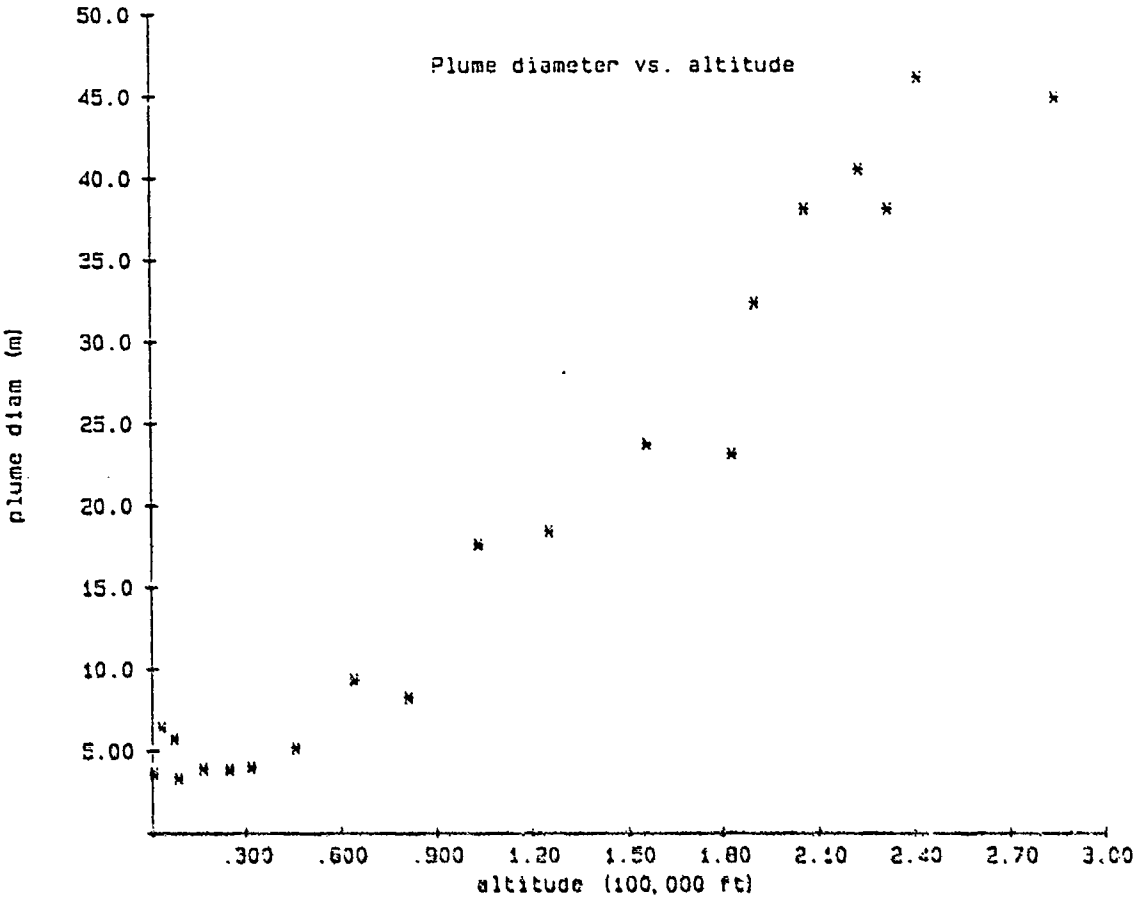


Figure 3. Plume diameter as measured by minor axis of the 50% brightness contour, accounting for the angular field of view per pixel and the distance to the rocket.

Since the contrail represents mostly solar scattering, photography of a night launch in the absence of solar scatter will eliminate the contrail problems.

The plume diameter and length are expected to increase with altitude in proportion to the mean free path of molecules in the earth's atmosphere. Since the mean free path varies inversely with density which varies negative exponentially with altitude, the plume diameter would be expected to vary exponentially with altitude. Plotting the logarithm of the plume diameter vs. altitude yields an approximately linear plot (Fig. 4) for altitudes below 100,000 feet in agreement with an exponential function.

B. Plume Brightness

The plume brightness was obtained from the axial plot of each plume image. The sky background was subtracted from each image to obtain the net signal from the plume. This assumes an optically thin plume, i. e. the sky is "seen" through the plume and the total brightness at the detector consists of the source plus the sky background. The results of this extraction is shown in Fig. 5. The brightness peaks at about 10 km, then falls back to about 2% of the maximum brightness due primarily to atmospheric absorption. The low brightness at about 5 km altitude (the anomaly in Fig. 5) is attributed to detector saturation.

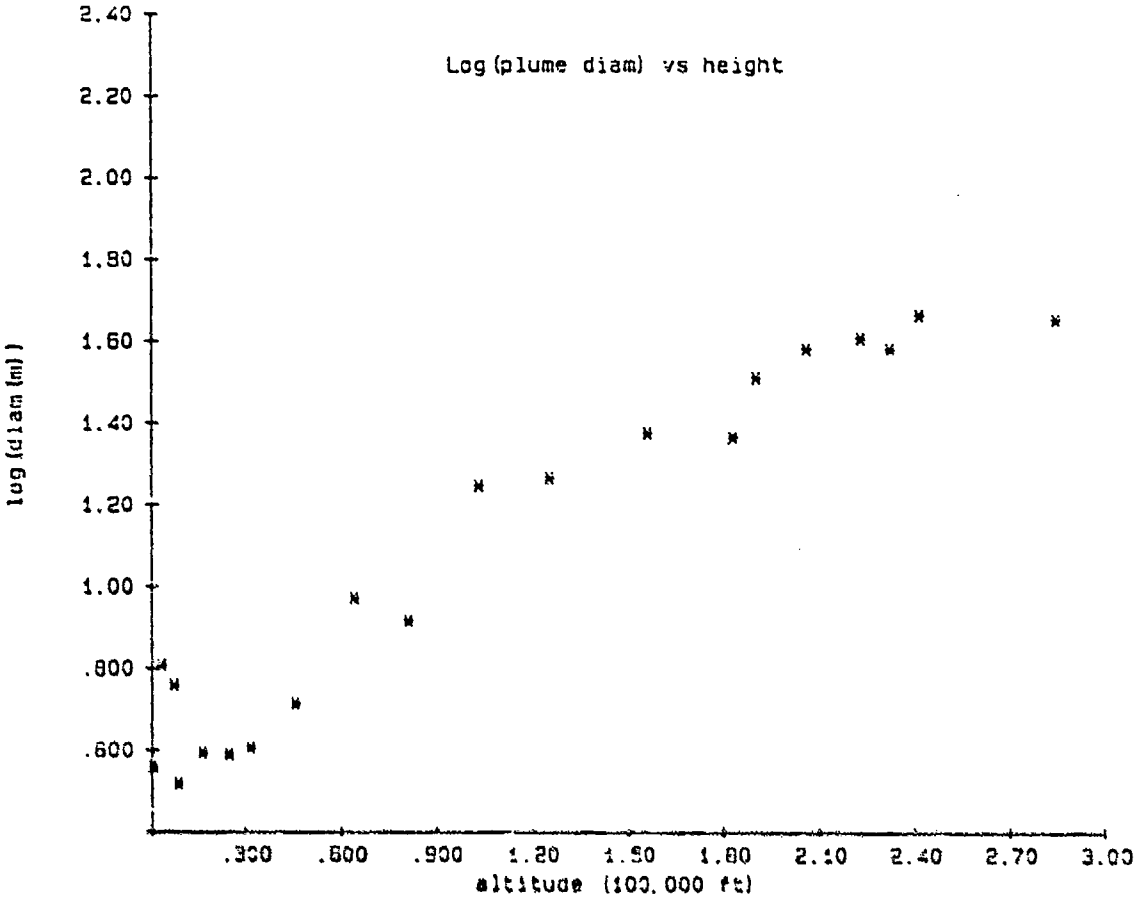


Figure 4. Logarithm of the plume diameter vs height.

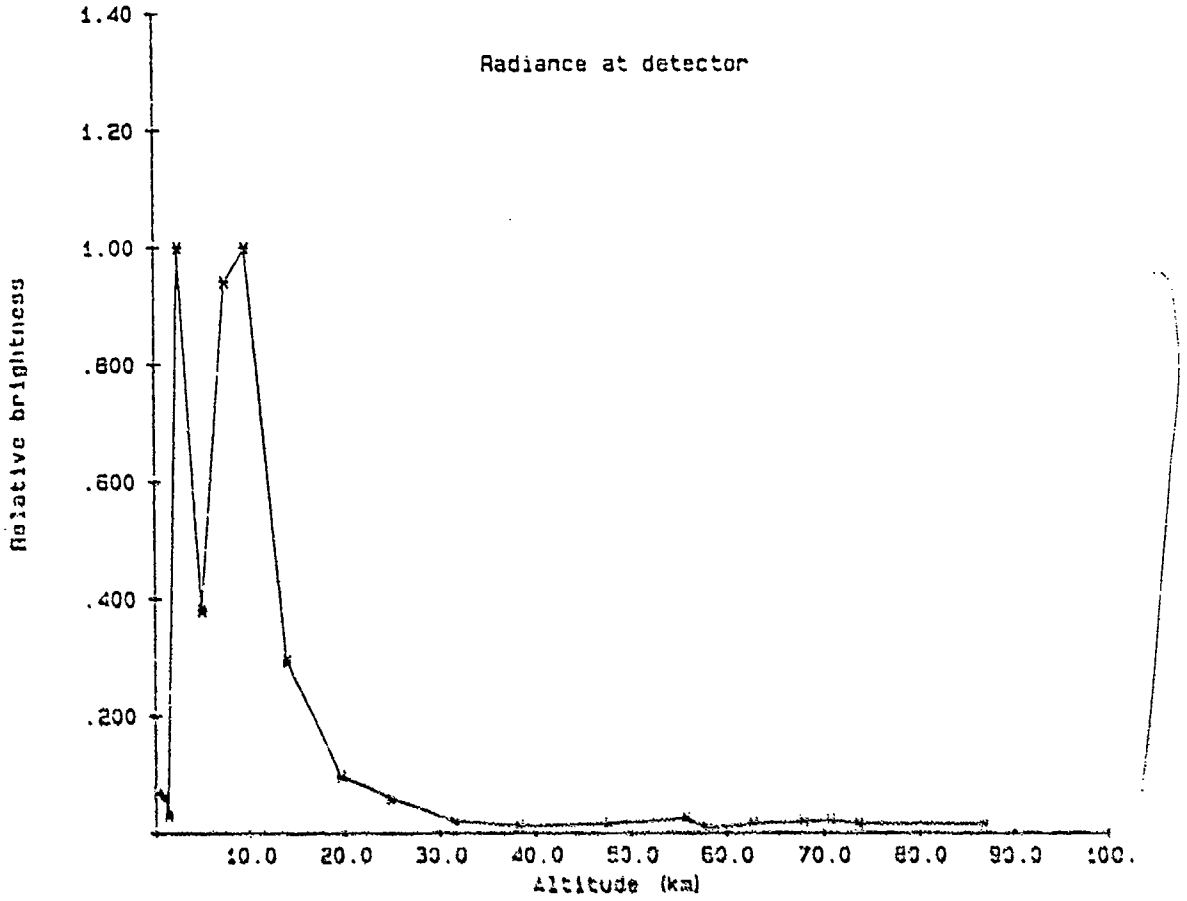


Figure 5. The relative brightness of the rocket plume seen at the detector assuming an optically thin plume.

This frame was exposed for 3 ms using the cathode gate on the camera, whereas the other frames in the neighborhood of the peak in brightness were exposed for 1 ms. From the author's work with the system, the AFGL UV imager saturates at about 0.5 volts which corresponds to 128 analog digital units (ADU). From the data the frame in question (3 ms exposure) shows a brightness of 126 ADU - clearly saturated.

The above discussion assumed a thin plume. The brightness was also extracted assuming an optically thick plume for all images. An optically thick plume is one through which the sky background cannot be seen. Instead of subtracting the sky background from the digital images, for a thick plume we subtract the dark current of the CCD detector in the solid-state video camera from the curves of Fig. 5. The dark signal is a constant (10 ADU) regardless of the cathode gate time because this is due to the dark current of the CCD during a single 1/30 sec frame interval. Fig. 6 shows the resulting brightness assuming an optically thick plume for all frames superimposed on the curve for an optically thin plume. For most of the images, however, except for the lowest altitude images, the optically thin plume is the best assumption regardless of the transmission of the plume itself. Much of the sky background is due to Raleigh scattering of sunlight and

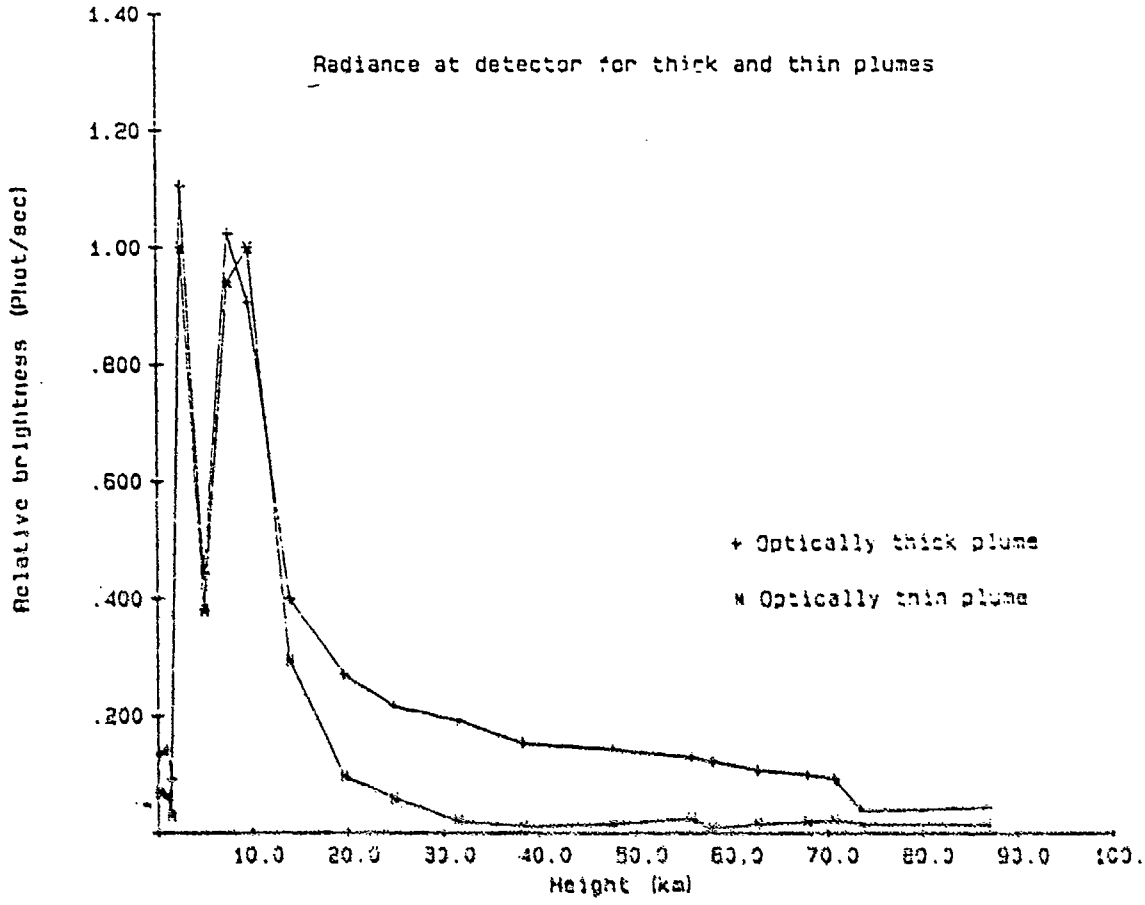


Figure 6. The relative brightness of the rocket plume seen at the detector assuming an optically thick plume compared with an optically thin plume.

most of the atmosphere producing the scattering lies between the source and detector, not beyond the source. Therefore the sky background is present of top of the plume. Even an optically thick plume does not obscure the sky background because the sky contribution arises between the source and detector. For the low altitude images, there is negligible difference in the relative brightness between thin and thick plumes because the plume is so much brighter than the sky background.

C. Atmospheric Transmission

In order to calculate the radiance at the source, the transmission of the atmosphere must be estimated. With photometry of an extended source, the absorbing medium between the source and the detector is the only factor which is distance dependent. Because brightness of an extended source is measured in power (or photons/sec) per area per solid angle, the brightness at a detector is independent of distance from the source if there is no absorbing medium. As distance increases, the smaller solid angle subtended by the camera aperture ($1/d^2$) is offset by the increased area of the source (d^2) seen by one resolution element. Thus the atmospheric transmission is the only factor to consider when calculating the brightness of the source from the brightness at the detector.

The transmission of the atmosphere is estimated by means of a simple Beer's law and exponential model of the atmosphere fitted to the estimated transmission at sea level. The Beer's law extinction coefficient, α, is defined as

$$dI/I = -\alpha(h)dh \qquad (3)$$

where $\alpha(h)$ is the extinction coefficient which depends on altitude h. $\alpha(h)$ is assumed to depend on the average density of the atmosphere which is an exponential function of altitude:

$$\alpha(h) = \alpha(0)\exp(-\tau h) \qquad (4)$$

where τ is the inverse scale-height = Mg/RT where R, T, M, and g are the ideal gas constant, temperature, molecular mass (kg/mole), and the acceleration due to gravity: all assumed constant for this approximation. For a slanted viewing direction through the atmosphere at an elevation angle θ, eq. (3) and eq. (4) become

$$dI/I = -\alpha(0)\exp(-\tau h)dh/\sin\theta \qquad (5)$$

which integrates to

$$I/I_o = \exp\left[\frac{\alpha(0)}{\tau \sin\theta}\left[\exp(-\tau h) - 1\right]\right] \qquad (6)$$

Eq. (6) is plotted in Fig. 7 for τ=8.5 km^{-1} and $\alpha(0)$ = .579 km^{-1}. The atmospheric transmission (I/Io) reaches a plateau for high altitudes as the source becomes above the bulk of the absorbing atmosphere.

For low altitudes the elevation angle of viewing is changing rapidly. Because the light path is essentially horizontal for the low altitudes and the total line of sight distance is fairly constant, the transmission of the atmosphere, I/Io, is assumed constant for elevations below 6 km. The extinction coefficient, $\alpha(0)$ is taken to be $.579 \times 10^{-1}$ km^{-1} which is calculated from the known values of aerosols, Rayleigh scattering, and O_3 at wavelength = 3100 A.

When the irradiance at the detector (Fig. 5) is divided by the transmission curve assumed for the atmosphere (Fig. 7) the radiance at the source is obtained and plotted in Fig. 8. The brightness reaches a maximum at about 10 km, then drops down and levels off. One would expect the brightness to diminish steadily with altitude because the plume steadily increases in size with altitude

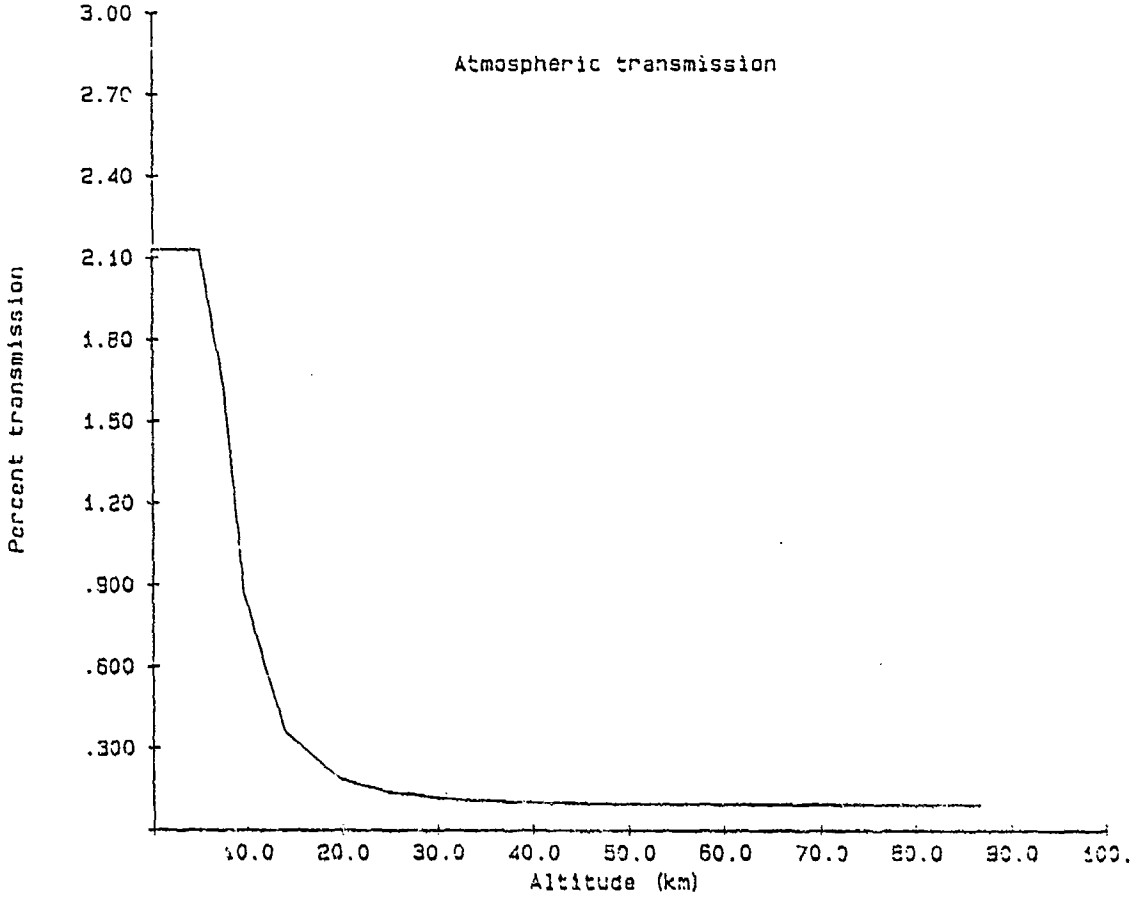


Figure 7. The transmission of the atmosphere through a 45 deg. path (eq.6) for the sea level extinction coefficient, $\alpha(0) = .579/km$.

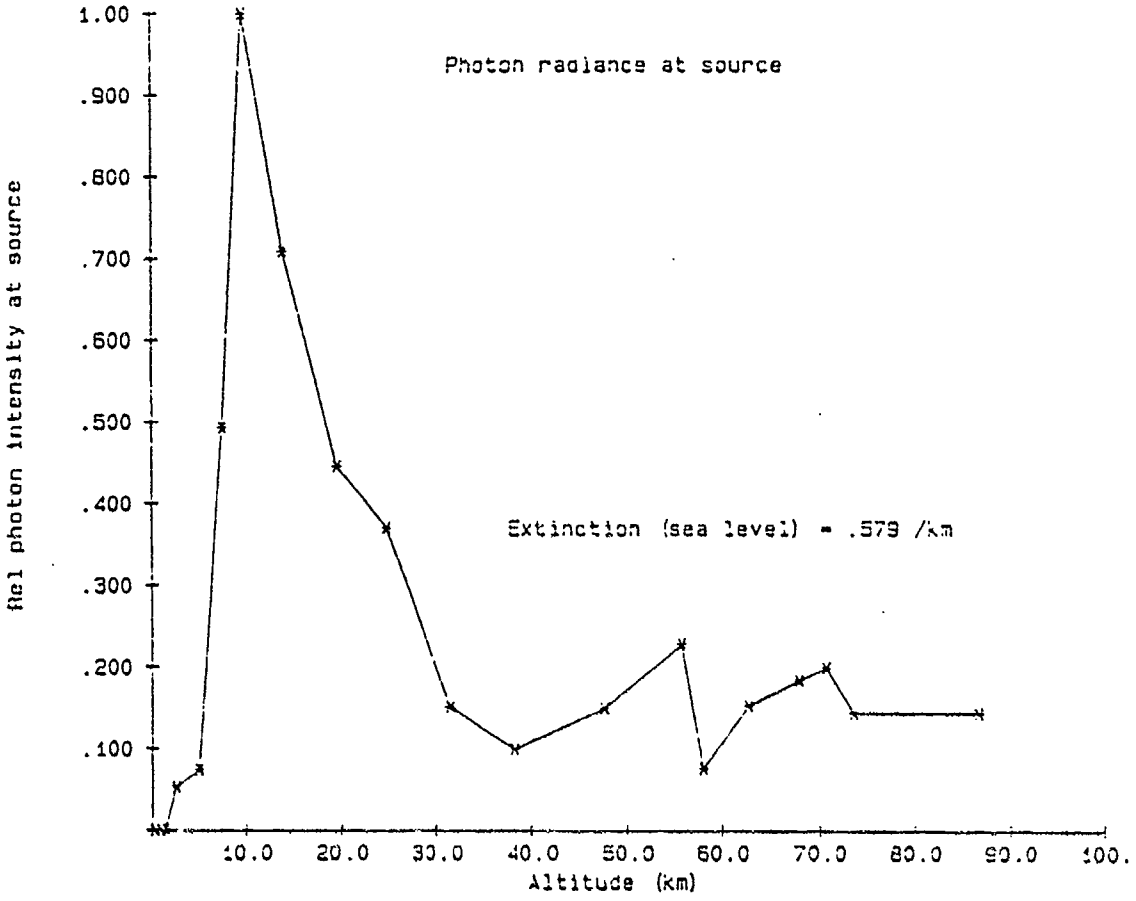


Figure 8. The source brightness (arbitrary units) of the rocket assuming the atmospheric transmission of Fig. 7: $\alpha(0) = .579$ /km.

with roughly a constant total energy, thus spreading the brightness over a larger volume. The discrepancy may be due to the lack of precision in the atmospheric transmission and the sea level extinction coefficient $\alpha(0)$. If $\alpha(0)$ is increased by a factor of 2 the source brightness curve (Fig. 8) changes dramatically to that in Fig. 9. Instead of a peak brightness at 10 km, the revised curve shows a constant brightness. Thus the estimated atmospheric transmission depends critically on the sea-level extinction coefficient which is a highly variable quantity dependent on local meteorological conditions. The simple exponential model for the atmosphere also has its limitations in the UV because different constituents which contribute significantly to the extinction coefficient have different scale heights and are also meteorologically dependent.

III. Molecular Diffusion

A. Description

One method to determine the mean free path of molecules in an ideal gas as well as the molecular sizes and particle density is to time the diffusion of a colored gas in air. This experiment is patterned after Alexandrakis (1978) and Rogers (1960). Bromine (or iodine) vapor is introduced into the bottom of a vessel,

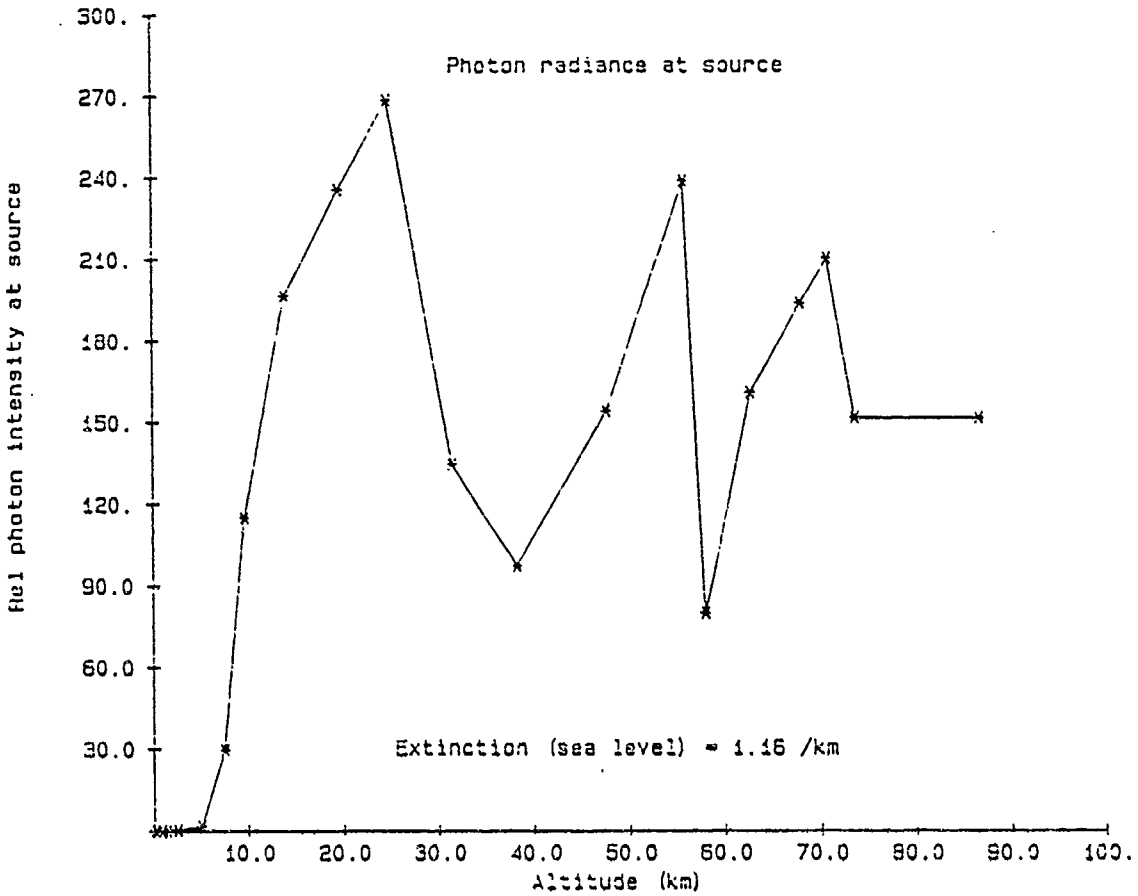
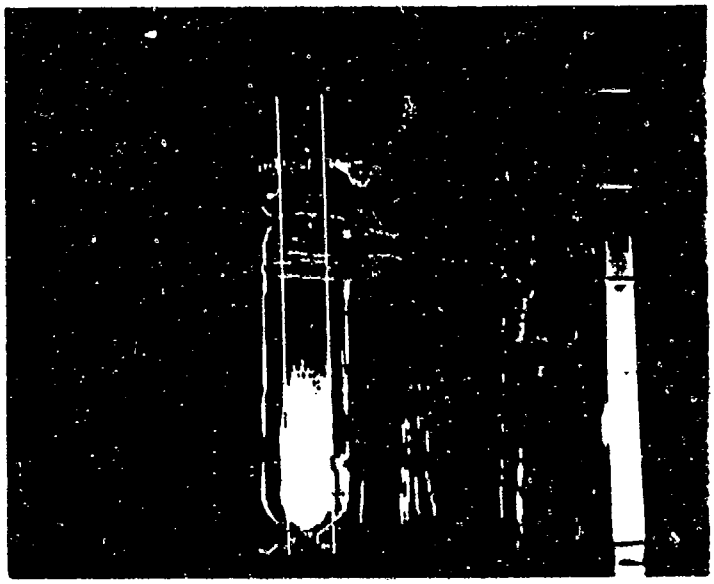


Figure 9. The source brightness similar to Fig. 7 but with sea-level extinction coefficient, $\alpha(0)$, twice as large (1.16 /km).

backlighted and photographed with the digital imaging system. As time progresses, the bromine vapor diffuses through the vessel as a Gaussian diffusion front (Feynman, et. al. 1963) and eventually becomes uniformly distributed throughout the container. In the traditional student experiment, the experimenter measures the time for the brown bromine (or purple iodine) cloud to reach a mark on the container. because of the Gaussian nature of a diffusion front, the student records the time when the mark is half as brown (or purple) as the base of the container. A photometer system greatly increases the precision of the diffusion front; an imaging photometer enables one to measure not only the time for diffusion, but also the spacial dependence of the bromine concentration.

Photograph 2 shows the apparatus photographed with the imaging hardware. The diffusion cell contains stopcocks both at the top and at the base. A spherical flask is attached to the bottom through the bottom stopcock. A few ml of liquic bromine is injected into the bottom flask and the bromine vapor is allowed to diffuse throughout the bottom flask with stopcocks closed. A syringe is fitted to the top stopcock with a piece of tubing. To begin the experiment, both stopcocks are opened and the syringe withdraws about 15 ml of air from the top stopcock. Because both stopcocks are open, the partial vacuum draws



Photograph 2. Digital image of the Bromine diffusion experiment. This is a difference image between the plain apparatus and the apparatus 1 minute after the admission of bromine vapor in the base. The highlighted columns indicate the region of interest in the later analysis.

bromine-laden air from the reservoir below. Care must be taken not to withdraw the bromine vapor too fast, otherwise turbulent mixing will take place in the diffusion cell and obscure the diffusion effect. The stopcocks are then closed and the cell photographed with the digital system about once every 30 sec as the bromine diffuses through the container. The bromine cell is backlighted with a mercury lamp and diffusing screen (no filter). A reference photograph is taken before the bromine is admitted so that relative intensities I/I_o may be measured. The relative concentration of bromine as a function of position is calculated from Beer's law:

$$[Br_2] = -\ln(I/I_o) \qquad (7)$$

Fig. 10 shows the Gaussian behavior of the bromine concentration at 135 sec. The solid curve represents a Gaussian function

$$\exp[-h^2/(2\,S_{rms}^2)] \qquad (8)$$

with S_{rms} = 6.12 cm. This mean diffusion length, S_{rms} is obtained by performing a quadratic polynomial regression on the $\ln[Br_2]$ vs height. S_{rms} is found from the square-root of the negative reciprocal of twice the quadratic coefficient. Thus the mean diffusion length S_{rms} is measured from the whole image - not the human eyeball photometer or a single channel photometer in one part of the diffusion cell.

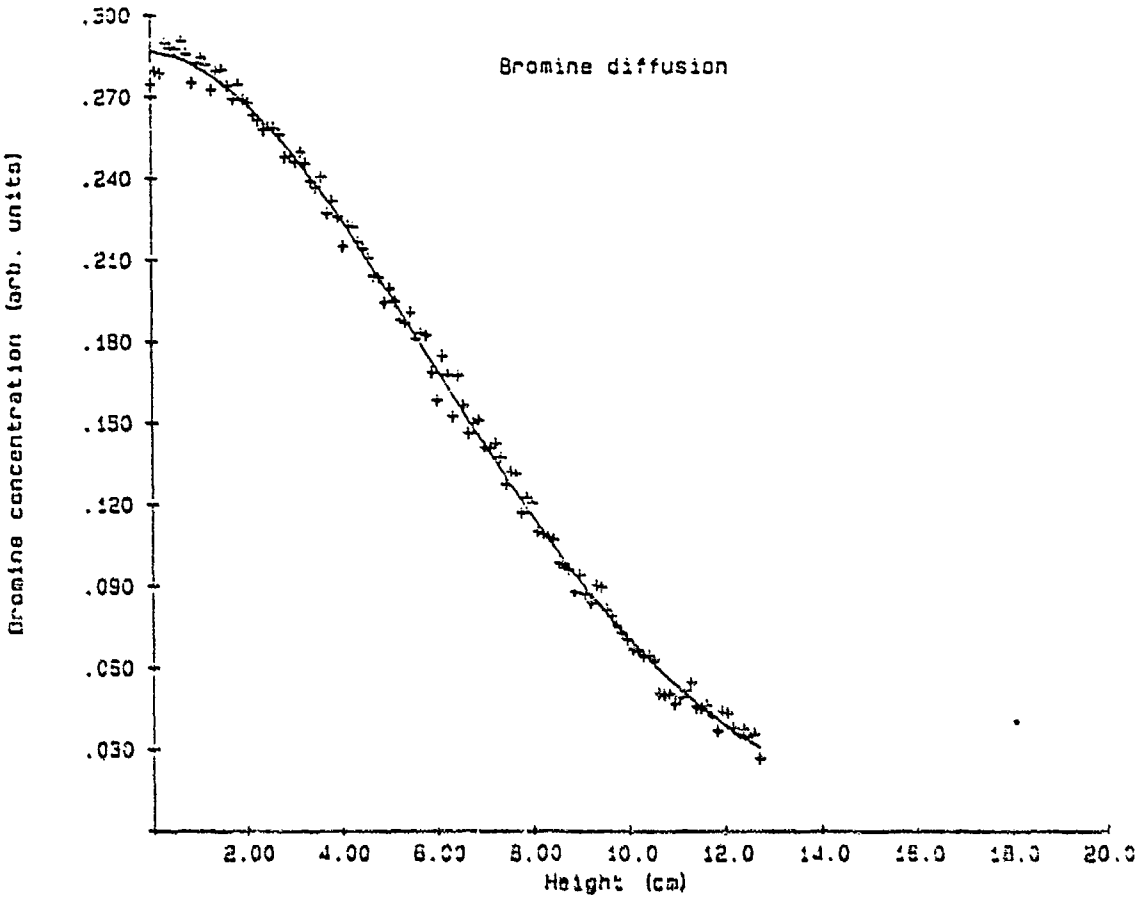


Figure 10. The concentration of bromine vapor as a function of height showing the Gaussian nature of the diffusion front. The solid curve is the fitted diffusion using the mean diffusion length S_{rms} as the parameter.

B. Theory

An excellent discussion of the theory of molecular diffusion based on a random walk model is found in Alexandrakis (1978). The mean free path ϵ is related to the mean diffusion distance, S_{rms}, by

$$\epsilon = \frac{S_{rms}^2}{\sqrt{(3P/\delta)}\ t} \qquad (9)$$

where P is the pressure, δ the density of bromine, and t the time. P is the total pressure (1 atmosphere, or 1×10^5 Pascal) in the diffusion chamber, not the partial pressure of bromine. Similarly, δ is the density of bromine vapor at the pressure and temperature of the experiment assuming the cell to be pure bromine and an ideal gas. The mean diffusion distance S_{rms} is obtained from the Gaussian diffusion front: the distance at which the bromine concentration is reduced by half. The mean free path ϵ depends on the molecular density n and the molecular diameter d as

$$\epsilon = 1/(n\pi d^2). \qquad (10)$$

The molecular density n is also related to the relative volumes of gas to liquid.

$$E_{vol} = D^3/d^3 \qquad (11)$$

where $D^3 = 1/(n)$ is the mean distance between molecules in

the vapor. Combining eqs. (10) and (11) along with kinetic theory for an ideal gas gives

$$d = \frac{S_{rms}^2 \pi}{(3P/\delta)^{1/2} E_{mps} t} . \qquad (12)$$

Here, δ is the density of pure bromine at the pressure and temperature of the experiment. Finally, the molecular density and Avogadro's number may be obtained by further algebraic manipulation:

$$N_a = \frac{(3P/\delta)^{3/2} E_{mps}^2 t^3 (Vol/mole)}{S_{rms}^o \pi^3} . \qquad (13)$$

Note that eqs. (10), (12), and (13) for the mean free path ℓ, the molecular diameter d, and Avogadro's number N_a are all based on easily measured experimental quantities: P, δ, E_{mps}, t, (Vol/mole), and the diffusion length S_{rms}. The S_{rms} is measured with the digital image photography.

C. Results

From the determination of S_{rms} to fit the data of Fig. 10, (6.12 cm at 135 sec), student laboratory measurements of δ (weighing a volume of air and vacuum and correcting for the difference in relative molecular masses of air and bromine), E_{mps} (by measuring the relative volume of liquid nitrogen after evaporation), and 22.4 x 10^{-3} m^3/mole, eqs (10), (12), and (13) give

$$\ell = 1.24 \times 10^{-7} \text{ m}$$

$$d = 7.08 \times 10^{-10} \text{ m}$$

and

$$N_A = 1.15 \times 10^{23} \text{/mole.}$$

These are reasonable numbers within an order of magnitude.

D. Problems and Improvements

The imaging photometer has enabled a reasonably accurate measurement of mean diffusion distance. However, because eq. (13) depends on the 6th power of $\bar{S}_{max}$, the determination of Avogadro's number can only be accurate to an order of magnitude. The factor of 6 difference from the accepted value could be accounted for by a factor of $\sigma_{1/e} = 1.5$ in $\bar{S}_{max}$. When the bromine is first admitted at time zero, unavoidable turbulence provides partial mixing of the bromine in air that the initial bromine distribution has a mean height of about 3 cm as opposed to zero cm assumed in the model. Modelling the diffusion from an initial distribution could be attempted to improve the experiment. Direct injection of bromine in the base of the diffusion cell with a syringe needle may improve the initial density gradient.

A word of caution: bromine is quite hazardous. A solution of Sodium Thiosulfate ($Na_2S_2O_3$) is prepared to

eliminate any bromine spills and to dissolve the bromine from the apparatus upon completion of the experiment.

IV. Critical Opalescence

The college possesses a demonstration device (Leybold Heraus Model 31740B1) to view a fluid near it's critical point. This apparatus consists of a cylindrical cell, filled with a critical amount of freon 115 which has a critical point (T_c, P_c) at 80 deg C and 30.8 bar. The cell is easily heated to above the critical temperature with steam from a water boiler. Below the critical temperature the fluid exists as two distinct phases (liquid and vapor) separated by a meniscus. Above the critical temperature the liquid phase boils so that the cell becomes totally single phase (vapor). As the cell is allowed to cool slowly in quasi equilibrium near the critical temperature, the fluid becomes increasingly turbid due to the critical density fluctuations. The large scattering of light from the critical density fluctuations is known as critical opalescence. When viewed in real color the cell takes on a pearly opalescence.

This cell permits photography of the transmitted light through end windows. For this experiment, a sodium discharge lamp backlights a diffuser to illuminate the

cell. The cell is photographed by the solid state video camera fitted with a close-up lens so that the 2.5 cm diameter cell nearly fills the video image. This experiment is being conducted by Carlos Benavente as a student research project still in progress. A search of the literature has yielded no previous photography of critical opalescence using a digital imaging system. Photographs 3, 4, and 5 show false-color digital images of the critical opalescence at $T > T_c$, $T \approx T_c$, and $T < T_c$.

From the transmission, I/I_o, the relative turbidity, τ, is calculated using Beer's law:

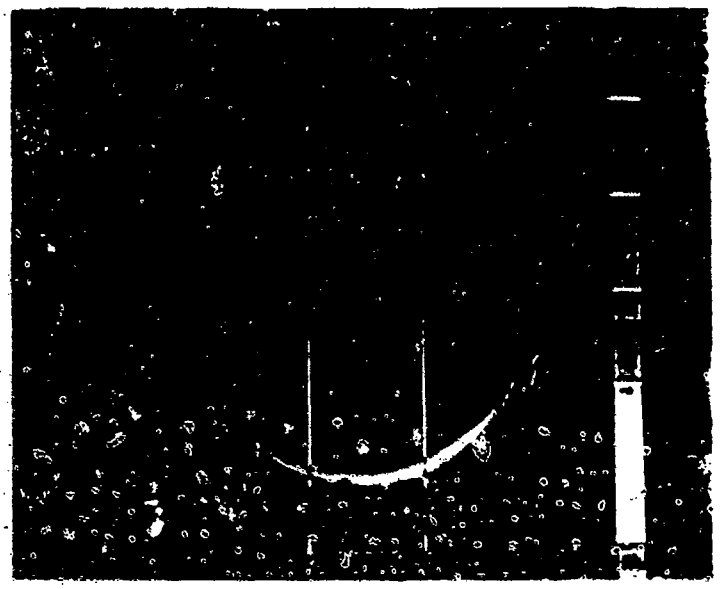
$$\tau = -\ln(I/I_o)$$

The turbidity is expected to diverge with temperature according to the power law:

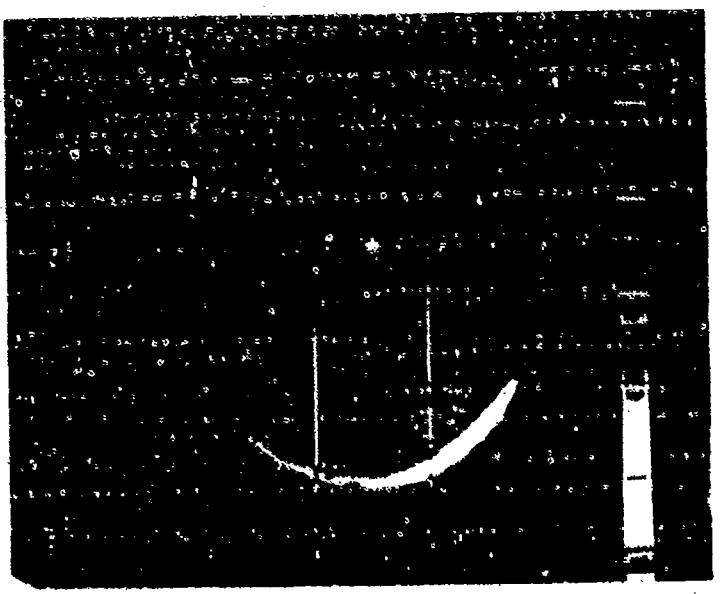
$$\tau = \left[\frac{T - T_c}{T_c}\right]^{-\alpha}$$

where α is known as a critical exponent (Stanley, 1971).

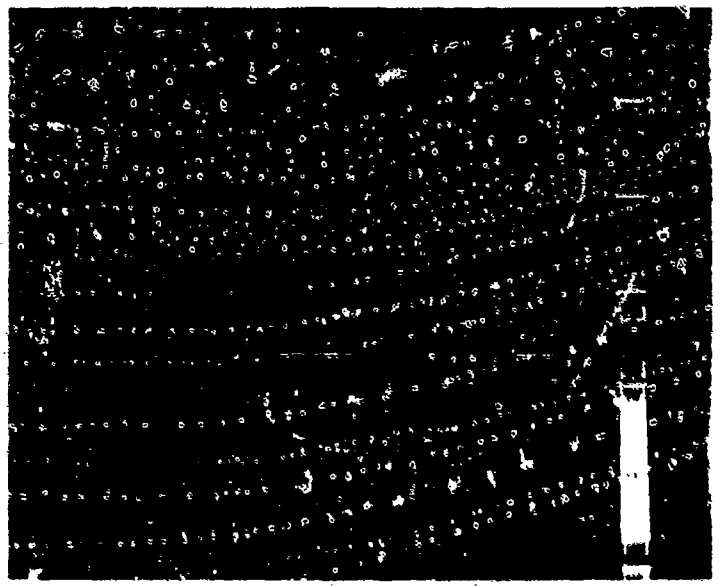
Because a fluid is so highly compressible near its critical point (the large compressibility gives rise to the large density fluctuations) it essentially sags under its own weight. Thus in the fluid cell near the critical temperature, the density varies from top to bottom in a continuous manner. By analyzing the image of the transmitted light vs. depth in the cell we are able to



Photogram 3. Digital image of critical coalescence (false color) for Freon 115 5 deg above the critical temperature.



Photogram 4. Digital image of critical coalescence (false color) for Freon 13 at the critical temperature.



Photograph 5. Digital image of critical opalescence (false color) for fresh 115 below the critical temperature showing the meniscus.

obtain the turbidity as a function of fluid density as well as temperature without having to adjust the volure of the cell. Fig. 11 is a preliminary plot of the turbicity vs. depth at 85 and 80 deg C.

V. Other Image Demonstrations

A number of image processing demonstrations were conducted to investigate the feasibility for further work. They are mentioned briefly below with the possibilities as tools for additional research.

A. Kármán Vortices

When a metal rod cooled to cryogenic temperatures is moved slowly through the air perpendicular to the rod's axis, the vapor clouds in the wake of the rod form vortex streets parallel to the length of the rod (see photographs 6 and 7). If the rod is stationary, a smooth curtain of condensed vapor is seen to fall from the rod. If the rod moves ≥ 1/2 m/sec, the smooth curtain breaks into the vortices known as Kármán vortices or streets (Rosnko, 1955). The image processing system represents a potential for further study of Kármán vortices - the shedding frequency vs. various parameters (rod diam, Reynolds number, rod speed, etc. . Photograph 7 shows that the vortices break off at an angle to the rod when the rod is rotated about an axis normal to the rod's length.

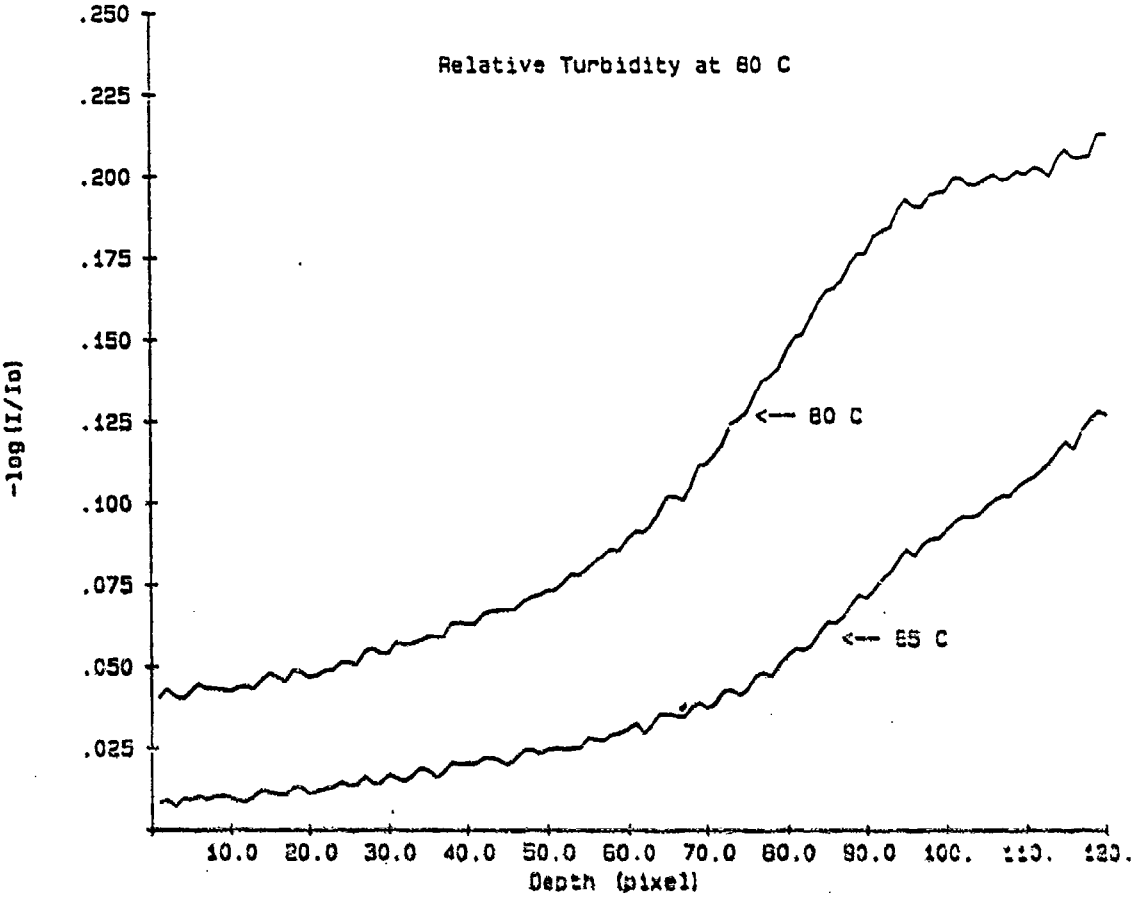
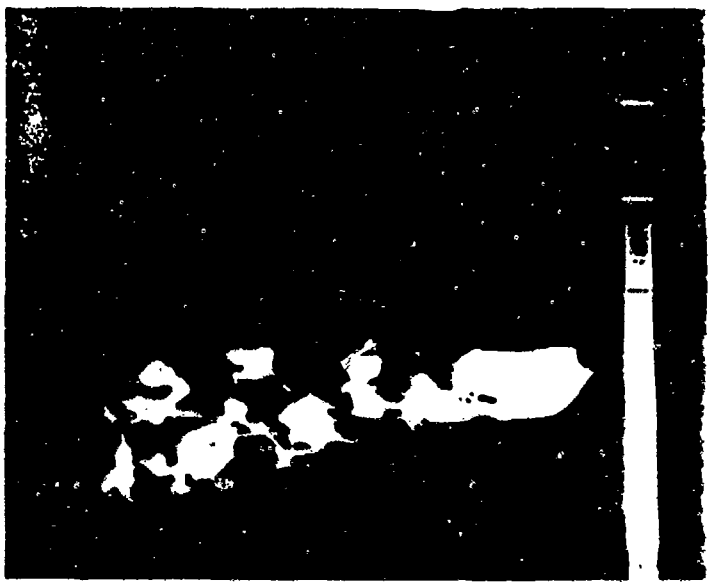


Figure 11. The turbidity of the critical opalescent cell vs. height at $T = T_c = 80$ C and $T = 85$ C.



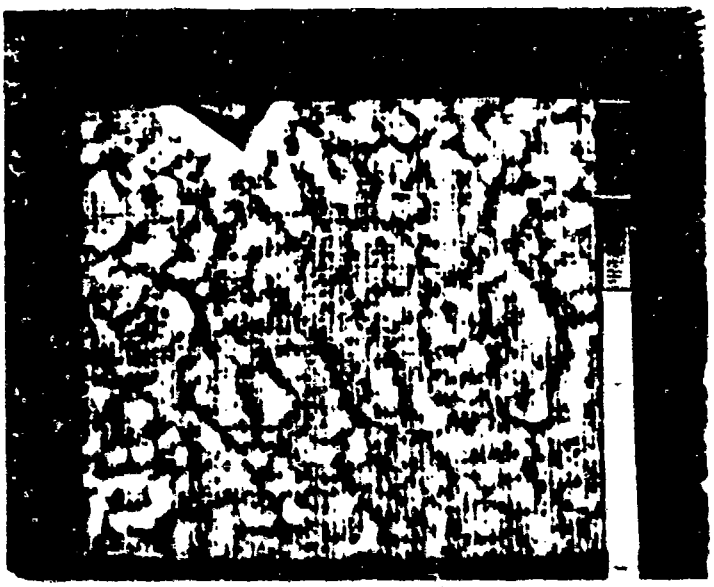
Photograph c. Digital image of Karman vortices.



Photograph 7. Digital image of Karman vortices viewed perpendicular to the rod. The rod is rotating about an axis normal to the picture.

B. Bénard cells

Convection of fluids are observed to flow in roughly polygonal cells, with the fluid rising in the central portion of the cell and descending along the cell boundaries. These can be seen in a cooling cup of coffee or tea, in a saucepan of heating water before boiling commences, and in cloud formations - often with the onset of a warm front. Photograph 8 shows the cloud formation after some image processing. Because the sun produces a large brightness gradient across the sky when Photograph 8 was taken, the image had to be "flattened" in order to see a uniform cloud pattern. The flattening was accomplished by subtracting a low-bandwidth (20 x 20 ave) filtered image from the original. The result (high spacial frequencies) was then contrast enhanced to make the cell-pattern visible. Photograph 9 shows the raw, but contrast enhanced, image before the flattening or high-pass filtering. Further study of Bénard cells should include the Reynolds number as a parameter especially in controllable "bench-top" experiments. The digital imaging system is a tool for making the convection cells visible.



Photograph 8. Digital image of Benard cell cloud formation. The brightness corner on top of the photograph is the corner of a building. The sun is further off the top of the picture.



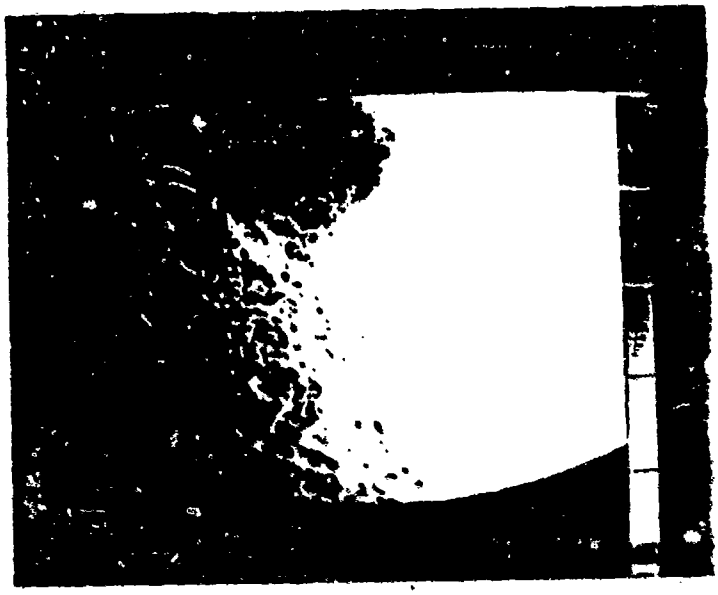
Photograph 9. Raw digital image of Benard cell cloud formations. The brightness gradient due to varying angular distances from the sun make the cloud structure nearly invisible.

C. Lunar Photography

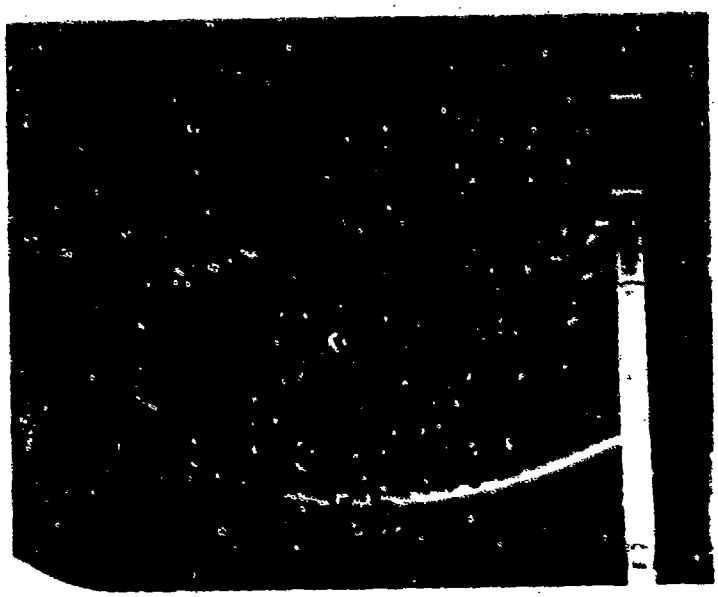
The video camera was mounted on the prime focus of a portable 8-inch Schmidt Cassegrain telescope and the moon was photographed (Photographs 10 and 11). The same type of flattening was applied to these photos as was done for the cloud formations. Further applications with lunar photography include relative brightness of differing regions, feature morphology, size measurements with the aid of a graphics cursor, and possibly searching for new craters, lunar librations, lunar distance through apparent size. These lunar observations are old and "passé" as far as original research is concerned. However, the image processing tools permit these experiments to be done by undergraduates at a small college.

D. Low Light "Time-exposure" Photography

Because many applications of an imaging system involve low-light illumination, and low-light image intensifiers are beyond the scope of the budget for this project, experiments at simulating time exposures by co-adding successive digital frames were performed. The PC-1024 frame grabber has the ability for real-time processing. This includes frame averaging at roughly 1/4 the real time rate. Thus a 1/2 sec. equivalent exposure should average 15 frames, but the duration of the exposure to obtain the



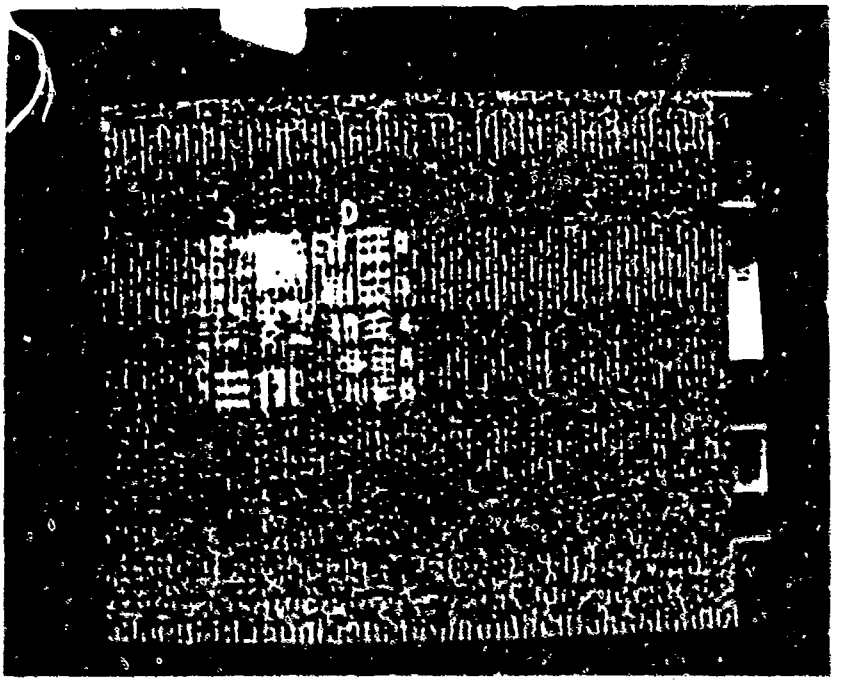
Photograph 13. Digital image of the moon through an 8-inch Schmidt-Cassegrain (Celestron-8) telescope. The central feature on the moon is the crater Tycho. This is a single frame (1/30 sec.) exposure.



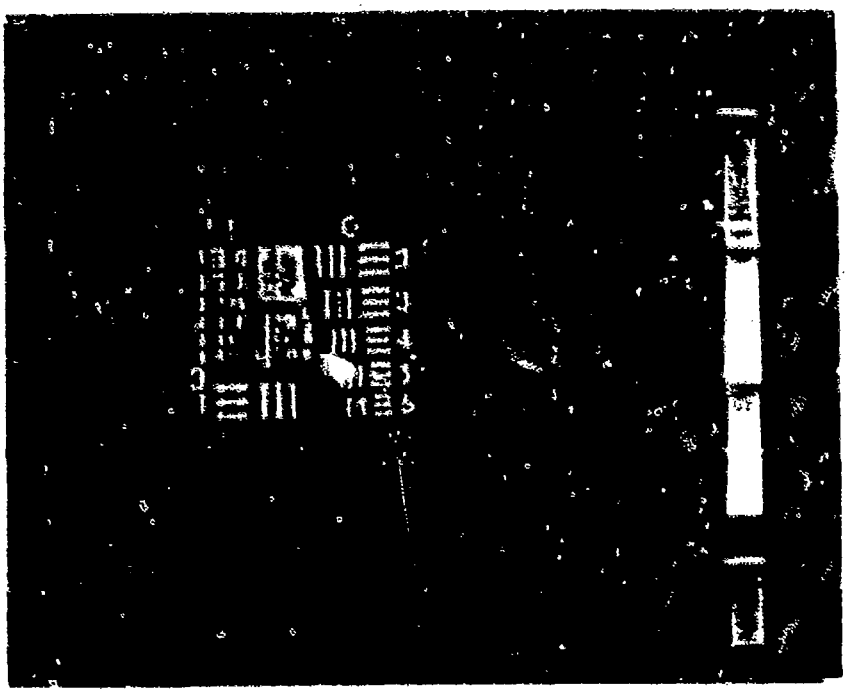
Photograph 14. Flattened high dynamic-range image made from photograph 13. The small features in the bright area around the planiform become more visible.

equivalent 1/2 sec of real-time exposure is about 2 sec, or four times longer.

To test the time exposure capabilities of the system, a transparent resolution chart was backlighted with an incandescent lamp and diffuser and photographed with the system. The lamp brightness and the CCD camera aperture were adjusted so that the live image on a TV monitor was barely discernable. When stretching the contrast in order to make the faint image visible, a Moiré effect between the digitizing rate (512 per line) and the CCD pixel pattern (380 per line) masks any image information. A conventional research quality CCD imaging system digitizes every CCD pixel at readout. Most of the Moiré effect can be cancelled, however, by grabbing a dark frame (with a lens cap on the camera) and subtracting. Photograph 12 shows the result of the net signal strenhed in contrast by a factor of 16. Photograph 13 shows the same image with the same illumination, but by averaging 48 video frames. The equivalent dark image (48 frame ave) was subtracted, but the contrast was stretched by a factor of 2. The improvement in signal/noise is dramatic. A single column of pixels through the group 0 horizontal bars in both images (Photograph 12 and 13) yields the pixel signal shown in Fig. 12 and Fig. 13. The signal/noise for the 48 frame average increased by a factor of 7/1.6 = 4.4. The central



Photograph 12. Low-light photograph of test pattern (single frame). Dark frame was subtracted and contrast stretched by a factor of 18. This image was near the threshold in the live TV display.



Photograph 13. Low-light photograph with same illumination as Photograph 12, but by averaging 48 frames and subtracting the dark image.

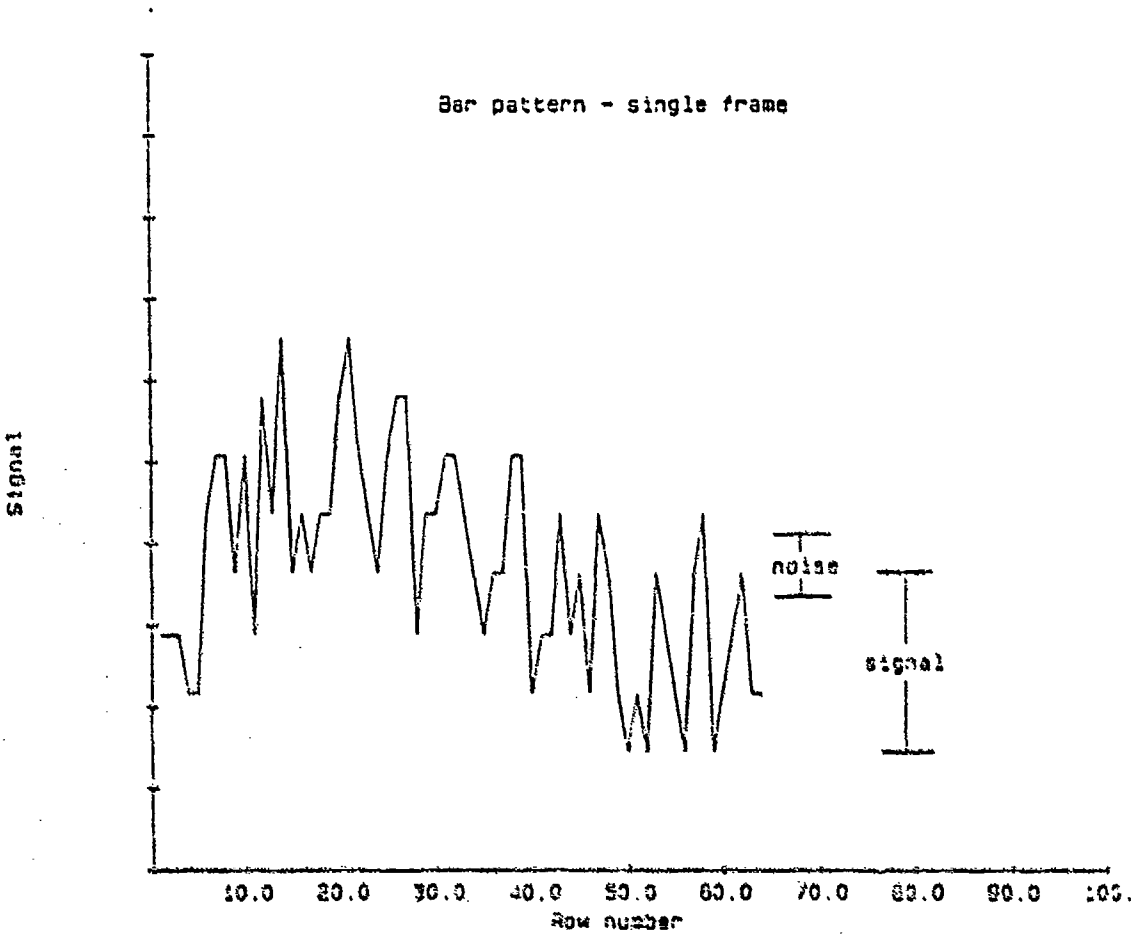


Figure 12. The pixel values from a single column of pixels from Photograph 12 - a single frame low-light image. The measured signal/noise is 40/25 = 1.6.

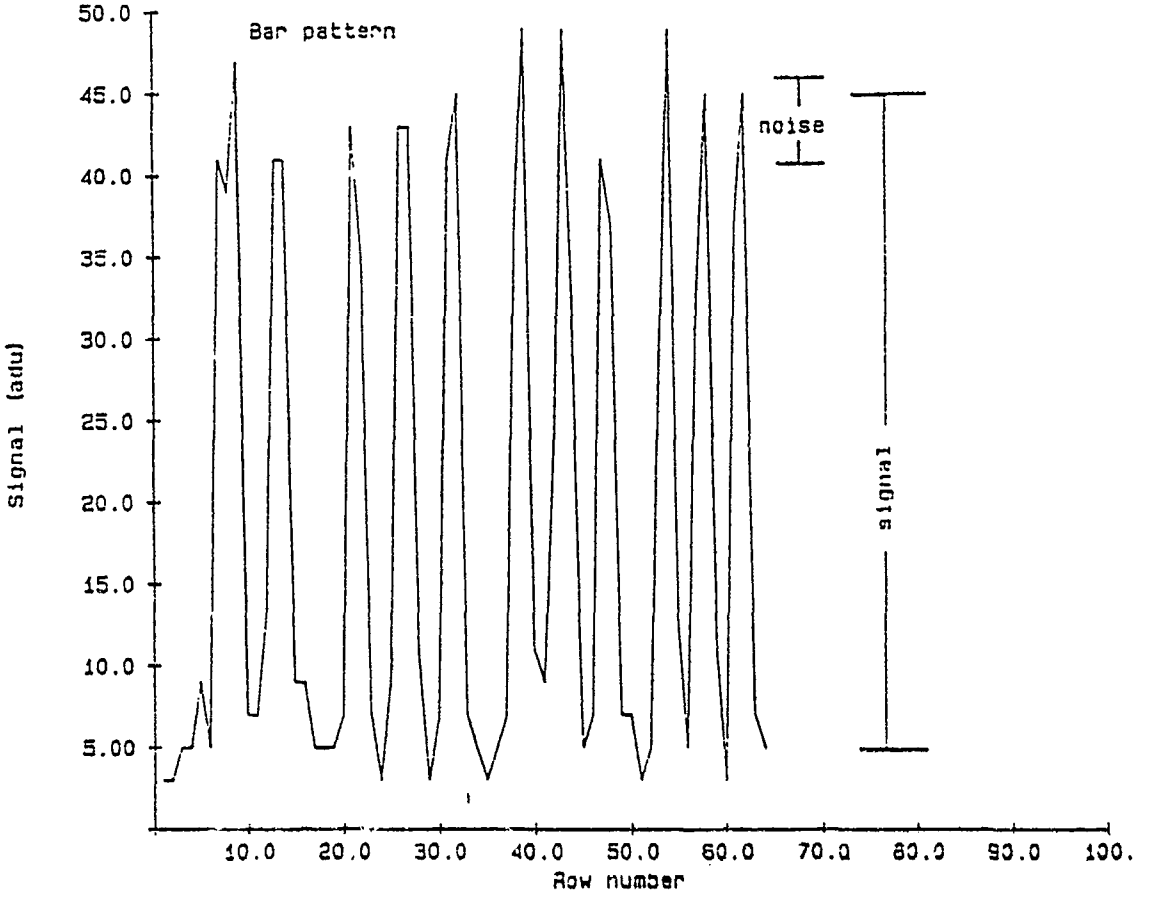


Figure 13. The pixel values from a single column of pixels from Photograph 13 - 48 frame average. The measured signal/noise is 40/6 ≈ 7.

limit theorem predicts $\sqrt{48} \approx 7$ for the signal/noise enhancement.

The resolution of the two images (single frame and 48 frame) from reading the line pairs groupings is increased only by a factor of two in contrast to the large enhancement in signal/noise. This demonstrates the tremendous subtle filtering and edge-detection capability in the human visual system. The visual resolution is maintained even for a noisy image (Niblack, 1985).

E. Linearity of Video Cameras

Linearity testing of video cameras was limited to focusing on a backlighted pinhole and adjusting the camera f-stop to control the illuminance on the detector. A miniature lamp was powered with a current controlled power supply at about 6 Volt to provide the backlighting. A diffusing screen ensured even illumination of the pinhole. The image of the pinhole was digitized and the average ADU reading for the spot recorded. An old Sony video camera (vidicon detector) and the Fairchild CCD camera (CCD 300E) were tested with this method and the results displayed in Fig. 14 and Fig. 15. The improved dynamic range of the CCD detector is obvious in the two figures. The highest point in the CCD curve saturates the 6-bit digitizer. Controlling the illuminance of the detector by means of the aperture stop on the camera lens is not the most precise

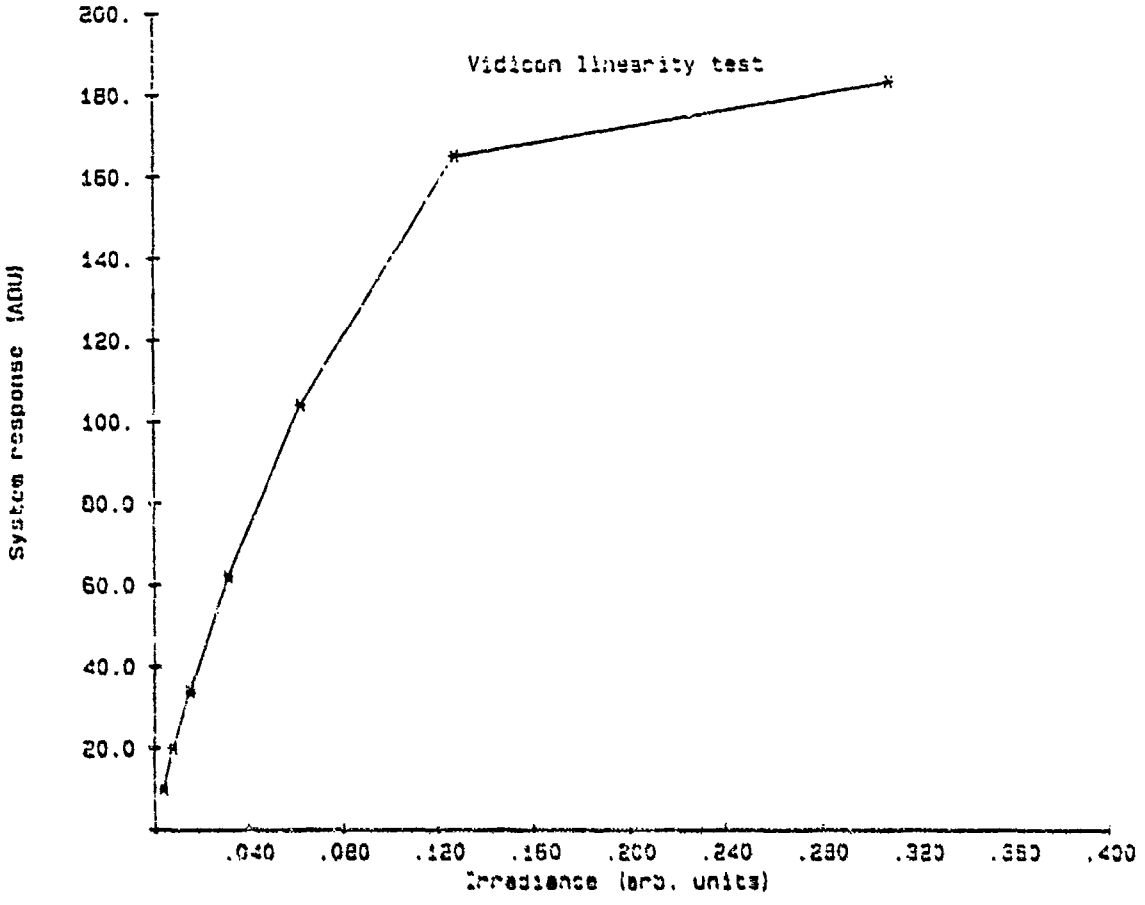


Figure 14. The video response of Sony video camera (vidicon). Notice that the camera detector saturates before the saturation of the 8-bit frame digitizer. The irradiance axis is $1/f^2$.

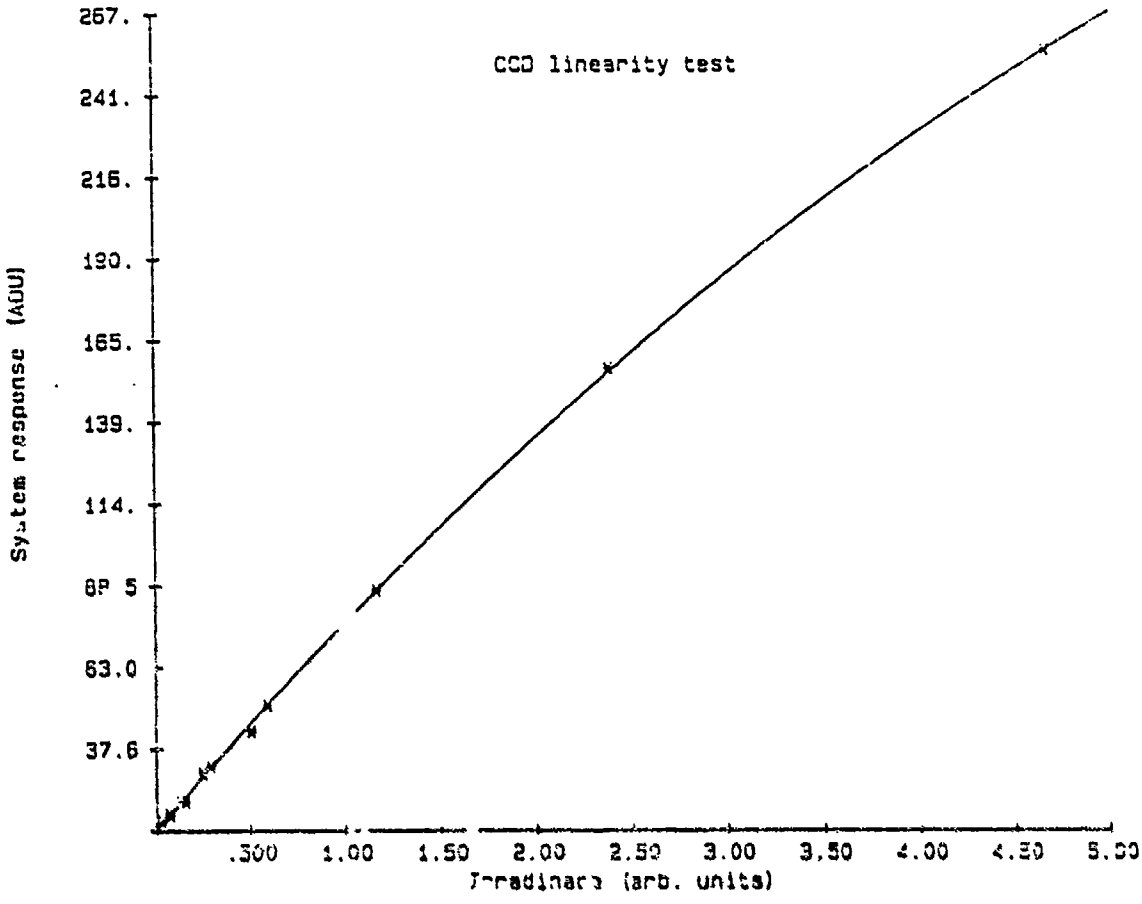


Figure 15. The video response of Fairchild (CCD 3002) video camera. The irradiance axis is 1/f2.

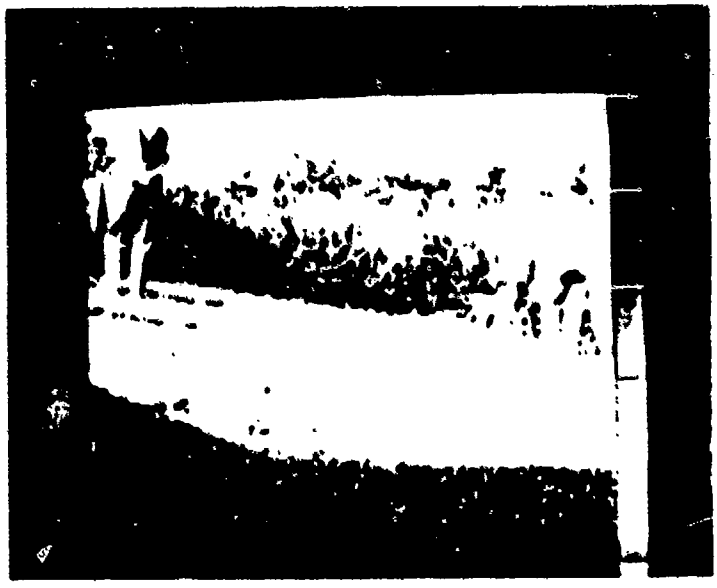
method, but the only method available. Since the illuminance varies as $1/f^2$, the illuminance values are not uniformly spaced. When the camera lens is focused on a close-up target, the effective aperture differs from the numbers printed on the lens barrel.

F. Differencing of "Busy" images.

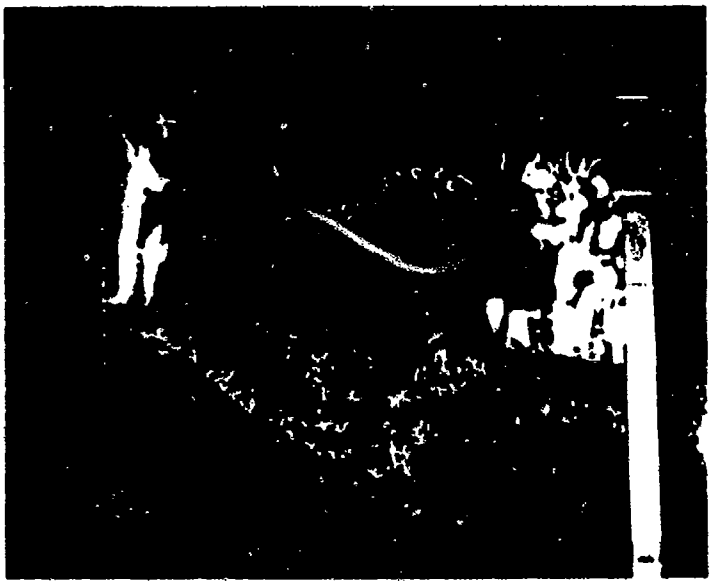
Subtracting similar images is one of the simplest image processing procedures. This has the power to look for small changes in the two images. This was demonstrated by photographing an area of grass, walking across the grass, then photographing the scene again. The two scenes look identical (Photographs 14 and 15), however the grass and leaves which were trampled, do not fully recover their original orientation, hence reflect light in a slightly different way. As soon as the images are subtracted the person's tracks show up (Photograph 16). An example application might be the comparison of astronomical images to search for planets, comets, variable stars and supernovae. Any real applications of image differencing will require images which can be correlated. In these demonstrations, the two images were obtained in rapid succession without disturbing the camera. If the camera orientation is not identical, or the average illumination is different, the two images must be cross-correlated (a computationally expensive procedure).



Photograph 14. Busy image of grass.



Photograph 15. 2nd busy image of grass after person walked on the grass.



Photograph 15. Difference between Photograph 15 and Photograph 14 showing the footprints in the grass.

G. Spark Discharges

One of the main goals of this proposal was to study spark discharges in the laboratory photographed with the imaging system. This proved impractical in the time available. Some sparks were recorded on video tape. The main problem is the lack of a suitable trigger for grabbing an image since a spark duration is less than 1 msec and is wholly contained on one video frame. Examining the video frames in still frame mode indicate considerable structure in the spark, but sync-tearing prevents satisfactory frame-grabbing in the still-frame mode. Use of a vidicon as opposed to CCD camera might help in the spark photography due to the persistence of vidicon tubes.

VI. Summary

The image digitizing and analysis system has successfully been used to analyze the rocket plume images and the procedures and system are available for future analysis. The system has successfully proven capable of analysis of molecular diffusion. Critical coalescence is also being investigated as a student project - an application of image analysis which has never been used before. A collection of other experiments have been demonstrated which show possibilities for further work.

Bibliography

Alexandrakis, George C., 1978, **American Journal of Physics 46**, 810-812.

Feynman, Richard P., Robert B. Leighton, and Matthew Sands, 1960, **The Feynman Lectures on Physics**, Addison Wesley, ch.6.

Niblack, W., 1985, **An Introduction to Digital Image Processing**, Prentice Hall.

Rogers, Eric M., 1960, **Physics for the Inquiring Mind**, Princeton Univ. Press p. 364.

Roshko, A., 1955, "On the Wake and Drag of Bluff Bodies", **Journ. Aeronaut. Sci. 22**(2), pp. 124-132.

FINAL REPORT NUMBER 15
REPORT NOT RECEIVED IN TIME
WILL BE PROVIDED WHEN AVAILABLE
Dr. William Cooper
760-6MG-081

1986 USAF-UES Mini-Grant

Research Program

Sponsored by the

Air Force Office of Scientific Research

Conducted by the Universal Energy Systems, Inc.

Final Report

Relationship Between Stages of Motor Learning and Kinesthetic Sensitivity

Prepared by:	Richard H. Cox
Academic Rank:	Professor
Deparment and University	Center for Human Movement Performance Kansas State University
Research Location:	Motor Learning Laboratory Center for Human Movement Performance Kansas State University
Date:	April 1, 1988
Contract No:	F49620-85-C-0013/SB5851-0360

Relationship Between Stages of Motor Learning and Kinesthetic Sensitivity

By

Richard H. Cox, Ph.D.

ABSTRACT

The relationship between three different measures of kinesthetic sensitivity and stages of learning a tracking and ballistic psychomotor task were studied. The three measures of kinesthesia included weight discrimination, linear positioning and angular positioning using either the dominant or nondominant hand. The two psychomotor tasks studied included the pursuit rotor tracking task and a discrete dart tossing task using either the dominant or nondominant hand. Little relationship exists between measures of kinesthesia regardless of which hand is used. Some evidence was observed that kinesthesia may be more important during the late as opposed to early stages of learning a tracking psychomotor task. However, when the motor task is ballistic as opposed to tracking, kinesthesia may be more important during the early or middle stages of learning as opposed to the later stages. Whether or not kinesthetic sensitivity is correlated with motor learning depends upon the nature of the task, the stage of learning and the technique used for measuring kinesthetic sensitivity. The findings of this research are so situation specific that it was difficult to arrive at a generalizable statement regarding the importance of kinesthesia to a psychomotor task such as flying an airplane.

Acknowlegements

I would like to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsoring my research. To be meaningful, scientific research must be conducted in an intellectually stimulating environment. The Human Resources Laboratory at Brooks AFB provided this environment. I would like to thank several members of the HRL staff, Major John Quebe, Dr. Jeffrey E. Kantor, Dr. Thomas R. Carretta, Dr. Frederick M. Siem, and Edison Watkins, for giving me the opportunity and the guidance necessary for my research.

I. Introduction

I received my Ph.D. from the University of Oregon in 1973 in the area of psychomotor learning and motor control. At Kansas State University I was a Professor in the Center for Human Movement Performance. It was while at Kansas State University that this research was initiated and completed with the assistance of the USAF-UES mini-grant program. I am currently Director of the School of Physical Education and a professor in the Psychomotor Division of the Human Performance Laboratory at Ball State University, Muncie, IN.

II. Objectives of the Research Effort

The Purpose of this research was to study the relationship between stages of motor learning and kinesthetic sensitivity. The psychomotor tests used in this investigation was the pursuit rotor task and a dart tossing task. Forearm kinesthetic sensitivity was measured through linear and angular positioning tasks, and through a weight discrimination task. Standard Fechnerian methods were utilized in calculating subject sensitivity.

III. Background and Literature Review

In a signal review of the literature related to cognitive and motor performance of Air Force pilots, Imhoff and Levine (1981) underscored the importance of kinesthetic sensitivity. They suggested that a test of kinesthetic sensitivity be included in any test battery disigned for pilot selection. As defined by Clark and Horch (1986), kinesthetic sensitivity refers to the awareness of the positions and movements of the limbs and other

body parts whether self-generated or externally imposed. Their definition excludes sensations arising from the visual and vestibular systems.

In combination with selected psychomotor tests of motor ability, kinesthetic sensitivity promises to be a powerful predictor of pilot performance and success. Correlations between motor performance and kinesthetic sensitivity have recently been observed to be as high as .40 and .43 (Fleishman & Rich, 1963; Cox & Walkuski, in press). Howver, the magnitude and strength of the relationship is related to the stages of motor learning. Fleishman and Rich (1963) examined the role of kinesthetic sensitivity in the learning of the two handed coordination task. In addition to taking the two hand coordination test, the 40 college students involved in the study were also tested for kinesthetic sensitivity through a weight discrimination task. The results of the study indicated that kinesthesia was significantly related to performance on the motor task in the late as opposed to early learning trials. Cox and Walkuski (in press) replicated Fleishman and Rich's basic design using the pursuit rotor task and three different measures of kinesthetic sensitivity. Results differed from those of Fleishman and Rich, in that Kinesthetic sensitivity as measured by an angular positioning task was observed to be significantly related to early but not late stages of learning.

From a theoretical perspective additional research is needed on the relationship between kineshesia and motor performance. Fleishman and Rich's results were explained in terms of Fitts and Posner's (1967) three stage theory of motor learning.

Conversely, Cox and Walkuski's results were explained in terms of Keele's (1973) motor programming model. The conflicting results could also be a function of differences in psychomotor tasks, methods of measuring kinesthesia and/or psychophysical methods employed. Regardless of the hypothetical reasons for the differences in results, it is quite obvious that continued scientific inquiry is needed and called for.

IV. Method

A. Sample - Subjects for this research were 91 male, right handed, college undergraduates. Informed consent was obtained from the subjects and confidentiality ensured.

B. Tests and Test Equipment - Two psychomotor learning tests and three forearm kinesthetic tests were administered to the 91 subjects.

1. Psychomotor Learning Tests

a. Pursuit Rotor Apparatus and Test - This novel continuous motor task was performed on a Lafayette model 30010 Pursuit Rotor set at 60 revolutions per min. A Lafayette Repeat Cycle timer was used to control the turning off and on of the rotary pursuit device. To begin the task, the subject stood in front of the pursuit rotor while holding the stylus in the left hand, resting the tip of the stylus in the center of the turntable. A "ready" signal was given and a cycle timer activated to initiate the task. Each of the 45 trials consisted of a 30 s tracking period followed by a 30 s rest period. During rest period subjects were instructed to place the tip of the stylus in

the center of the turntable. Time on target scores were recorded and knowledge of results given to the subjects during the rest interval (three to five seconds after trial completed). For analysis purposes, the 45 trials were broken down into 15 three trial blocks. Performance was recorded as time on target (TOT).

b. Dart Tossing Task - In this dynamic discrete motor learning task the subjects were required to toss 150 darts while standing behind a 2.4 m restraining line. The subjects tossed the darts in 5-blocks of 5-tosses (25 tosses) followed by a 3-min rest break. This process was repeated six times for a total of 150 tosses. Score values were recorded as deviations in half cm (5 mm) from the outside edge of the bull's-eye which had a radius of 1.75 cm. A dart landing in the bull's-eye was given a point value of zero for no deviation. A dart landing off of the target received a deviation value of 25 cm (250 mm). The diameter of the entire competition quality target was 45 cm. Matched competition quality darts were used for all tosses. For analysis purposes, the 30 5-toss trial blocks were broken down into 10 15-toss trial blocks. Performance was recorded to the nearest error (mm) score.

2. Tests of Kinesthetic sensitivity

a. Weight Discrimination Test - The weight discrimination task is a measure of kinesthetic sensitivity based on a subject's ability to discriminate among weights. Weighted

cannisters of identical size were used. The cannisters weighed 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, and 125 grams. The standard or comparison weight was the 100 gram cannister. For test administration, the subject was seated at a table wile wearing a blindfold. The subject rested his arm on a table while the weights were brought into the testing area. The subject was instructed to lift the presented weight from the wrist using the thumb, index and middle fingers. The weights were presented in pairs, the standard weight followed by a random presentation of the test weights. The experimenter placed and removed the weights from the subject's grasp, stating "standard" when the 100 g weight was presented and "test" when the second weight was presented. The subject's response of whether the second weight was heavier or lighter than the standard was recorded. There was a five second interval between the subject's response to the presentation of the second weight and the experimenter's presentation of the standard weight for the next comparison. No feedback was given. These procedures were repeated until the subject experienced each of the test weights 10 times for a total of 100 trials. From the data, using the method of constants, the difference limen (JND) was calculated as outlined in Galanter (1962).

b. Angular Positioning Device and Test - This test of kinesthetic sensitivity was performed on an angular positioning device. The device consisted of a hardwood box 20 in. high, 27 in. wide and 18.5 in. deep. In the

back of the box, a slot 10 in. high is cut. From the inside of the box, a "T" shaped metal handle extended through a frictionless pivot area to the outside front of the box and connected to a counter weight balanced pointer 33 cm long. A curvilinear scale measuring from zero to 180 deg was applied to the front of the box. The test consisted of 39 trials. Thirteen angular movements were presented three times. The angular movements used involved movements of 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, and 150 deg. The angles were presented randomly to the subject. The movement task required the subject to attempt to replicate the end point of an angular movement while concentrating on the movement location and extent. The subject was seated behind the device and asked to extend his hand through the slot in the back of the box, grasping the "T" shaped handle between the index and middle fingers with the palm facing up. With the pointer at the zero position on the scale, the required movement was pronation of the hand. The experimenter moved the external pointer, while the subject passively grasped the handle, to one of the criterion angles. After two seconds, the subject was instructed to release the handle while while the experimenter returned the pointer to the starting position. The subject was then asked to regrasp the handle and to actively attempt to replicate the end point of the criterion angle. Error was measured to the nearest degree. A 10 s time interval

took place between the end of a trial and the beginning of the next trial. This procedure was repeated until all 39 trials were completed. In analyzing the data, an error score was calculated using Henry's E (1975).

c. Linear Positioning Device and Test - The device consisted of a four inch metal rod mounted vertically on a frictionless slide which run on two stainless steel rods. A pointer was attached to the slide and a linear scale marked off from zero to 100 cm placed along the slide. The actual test consisted of 39 trials made up of 13 distances repeated three times. The distances used were: 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, and 80 cm. The distances were randomly presented. The subject's task was to attempt to replicate the end point of a passive linear movement. To begin the test, the subject was seated and blindfolded while the linear positioning apparatus was placed in front of him with the 30 cm mark at the center of the subject's upper torso. With guidance from the experimenter, the subject grasped the handle of the device loosely with his hand. With the pointer at the starting position (zero on the scale), the experimenter moved the subject's hand along the slide to one of the criterion positions. After two sec at the criterion position, the subject was instructed to release the handle. The experimenter then returned the slide to the starting postition. The subject was then asked to regrasp the handle and immediately attempt to replicate the previous end location of the movement. Error was measured

to the nearest centimeter. The subject was then asked to release the handle. A 10 sec time interval took place between the end of a trial and the beginning of the next trial. This procedure was repeated until all 39 trials were completed. For analysis purposes, an error score was calculated using Henry's E (1975).

C. Procedures and Design of Experiment

Laboratory testing consisted of three tests of kinesthetic sensitivity and two novel motor learning tasks. On day one the subjects were tested for hand dominance using a 14 item discriminative questionnaire developed by Crovitz and Zener (1962). Only subjects exhibiting clear right hand dominance were retained for further testing (score of 30 or less). On day one each subject's kinesthetic sensitivity was ascertained using the three tests of kinesthetic sensitivity. On day two each subject practiced the pursuit rotor task, and on day three the dart tossing task.

In terms of design, the first 50 subjects were assigned to a nondominant or left hand experiment while the next 41 subjects were assigned to a dominant or right hand experiment. The two parallel experiments were formed in order to compare learning when the task was truly novel (left hand) as opposed to novel (right hand). Thus, the subjects either used the dominant or nondominant hand for all motor learning and kinesthetic sensitivity tests.

D. Variables

All variables were continuous in nature, with the scores

from the kinesthetic sensitivity tests serving as the independent variables and the scores from the motor learning tasks serving as the dependent variables. The variables were analyzed using standard correlation techniques.

V. Results

A. Kinesthetic Sensitivity

1. Left Hand - The average kinesthetic sensitivity measure (jnd) for the 50 subjects on the weight discrimination task was 6.36 g with a standard deviation of 2.53 g. The average angular positioning sensitivity score was 8.38 deg with a standard deviation of 2.21 deg. Finally, the mean linear positioning kinesthetic sensitivity score was 3.28 cm with a standard deviation of 0.78 cm. A simple correlational analysis revealed that none of the measures of kinesthetic sensitivity were significantly related to each other at the .05 level.

2. Right Hand - The average kinesthetic sensitivity measure (jnd) for the 41 subjects on the weight discrimination task was 6.93 g with a standard deviation of 2.47 g. The average angular positioning sensitivity score was 7.80 deg with a standard deviation of 2.36 deg. Finally, the mean linear positioning kinesthetic sensitivity score was 3.40 cm with a standard deviation of 0.61 cm. A simple correlational analysis revealed that only angular and linear positioning were significantly related to each other, $r = 0.391$, $p < 0.05$.

B. Pursuit Rotor and Kinesthesia

1. Left Hand - The results of a repeated measures ANOVA on the

15 3-trial blocks revealed a significant learning effect across blocks. The nature of the learning curve is illustrated through mean performance scores in Table 1. As can be observed in Table 1, performance scores improved steadily across blocks 1 through 15. Further analyses, using the Duncan post hoc multiple comparison test, revealed that significant performance increments occurred among trial blocks one through 10 and that a clear asymtote was reached during trials 11 through 15.

Simple correlations between trial block pursuit rotor tracking scores and three measures of kinesthetic sensitivity are displayed in Table 1. As can be observed in this Table, a significant relationship does not exist between pursuit rotor learning and kinesthetic sensitivity when kinesthesia is measured through angular positioning or linear positioning. However, in the case of weight discrimination, a low but significant relationship exists between kinesthesia and pursuit rotor performance for trial blocks 14 and 15.

Insert Table 1 About Here

---

2. Right Hand - The results of a repeated measures ANOVA on the 15 3-trial blocks revealed a significant learning effect across blocks. The nature of the learning curve is illustrated through mean performance scores in Table 2. As can be observed in Table 2, performance scores improved

steadily across blocks 1 through 15. Further analyses, using the Duncan post hoc multiple comparison test, revealed that significant performance increments occurred among trial blocks one through 10 and that a clear asymtote was reached during trials 11 through 15.

Simple correlations between trial block pursuit rotor tracking scores and three measures of kinesthetic sensitivity are displayed in Table 2. As can by observed in this Table, a significant relationship does not exist between pursuit rotor learning and kinesthetic sensitivity when kinesthesia is measured through weight discrimination, angular positioning or linear positioning.

Insert Table 2 About Here

C. Dart Tossing and Kinesthesia

1. Left Hand - The results of a repeated measure ANOVA on 10 15-trial blocks revealed a significant learning effect across trial blocks. The nature of the learning curve is illustrated through mean performance scores in Table 3. As can be observed in Table 3, performance scores improved steadily across blocks. Further analyses, using the Duncan post hoc multiple comparison test, revealed that significant performance increments occurred among blocks one through three with an asymtote being reached following trial block number four.

Simple correlations between trial block dart tosses and

the three measures of kinesthetic sensitivity are displayed in Table 3. As can be observed in this Table, a significant relationship does not exist between dart tossing and kinesthetic sensitivity as measured by weight discrimination or angular positioning. In the case of linear positioning a significant relationship was observed between dart tossing and kinesthetic sensitivity during trial block one, r = 0.446.

Insert Table 3 About Here

---

2. Right Hand - The results of a repeated measure ANOVA on 10 15-trial blocks revealed a significant learning effect across trial blocks. The nature of the learning curve is illustrated through mean performance scores in Table 4. As can be observed in table 4, performance scores improved steadily across blocks. Further analyses, using the Duncan post hoc multiple comparison test, revealed that significant performance increments occurred among blocks one through three with an asymtote being reached following trial block number four.

Simple correlations between trial block dart tosses and the three measures of kinesthetic sensitivity are displayed in Table 4. As can be observed in this Table, a significant relationship does not exist between dart tossing and kinesthetic sensitivity as measured by weight discrimination or linear positioning. In the case of angular positioning a significant relationship was observed between dart tossing

and kinesthetic sensitivity during trial blocks three through six.

Insert Table 4 About Here

VI. Discussion

Relationships among measures of kinesthetic sensitivity are either nonexistent or minimal at best. Using the left or nondominant hand, none of the three measures of kinesthesia were significantly related to each other. The only significant relationship between the three measures of kinesthesia occurred between linear and angular positioning when the right or dominant hand was being used. However, even this relationship was minimal since the two measures only had 15.2 % of the variance in common, $r = 0.391$, $p < 0.05$.

Little or no relationship exists between measures of kinesthetic sensitivity and the rotary pursuit tracking task at any stage of motor learning. The only exception to this conclusion occurred when subjects used their nondominant (left) hand and kinesthesia was measured using the weight discrimination task. In this case significant correlations were observed between pursuit rotor tracking and kinesthetic sensitivity for trial blocks 14 and 15 ($r = -0.324$ and -0.317 respectively). The negative correlations were consistent with a positive relationship since a good pursuit rotor score was a high score while a good JND score was a low score. This finding was consistent with Fleishman and Rich's (1963) results since they

also observed a significant relationship between weight discrimination and the two hand coordination tracking task during the later stages of motor learning. In this respect the results of this research support the notion that kinesthetic sensitivity is most important during the later stages of learning a pursuit tracking motor skill (Fitts & Posner, 1967).

The results of this investigation differ markedly when the motor learning task is ballistic and discrete in nature. When the dart tossing task was performed with the nondominant (left) hand, a significant relationship between stage of learning and kinesthesia occurred in a very early stage of learning and when kinesthesia was measured using the linear positioning task. Conversely, when the task was performed with the dominant (right) hand, significant correlations were observed during the middle stages of learning and when kinesthesia was measured using the angular positioning task. Apparently, when the discrete ballisitic task is truly novel (left hand) kinesthetic sensitivity, as measured by linear positioning, is important during the initial or early stage of learning. This result would be consistent with Keele's (1973) motor program theory.

Whether or not kinesthetic sensitivity is correlated with motor learning depends upon the nature of the task, the stage of learning and the technique used for measuring kinesthetic sensitivity. The findings of this research are so situation specific that it is difficult to arrive at a generalizable statement regarding the importance of kinesthesia to a psychomotor task such as flying an airplane.

References

Clark, F.J., & Horch, K.W. (1986). Kinesthesia. In K.R. Boff, L. Kaufman & J.P. Thomas (Editors), Handbook of perception and human performance. New York: John Wiley and Sons.

Cox, R.H., & Walkuski, J.J. (in press). Kinesthetic sensitivity and stages of motor learning. Journal of Human Movement Studies.

Crovitz, H., & Zener, K.A. (1962). A group test for assessing hand and eye dominace. American Journal of Psychology, 75, 271-276.

Fitts, P.M., & Posner, M.I. (1967). Human performance. Belmont California: Brooks/Cole.

Fleishman, E.A., & Rich, S. (1963). (1963). Role of kinesthetic and spatial-visual abilities in perceptual motor learning. Journal of Experimental Psychology, 66 6-11.

Galanter, E. (1962). Contemporary psychophysics. In R. Brown, E. Galanter, E. H. Hess, & G. Mandler (Editors), New directions in psychology. New York: Holt, Rinehart and Winston.

Henry, F.H. (1975). Absolute error versus "E" in target accuracy. Journal of Motor Behavior, 7, 227-228.

Keele, S.W. (1973). Attention and human performance. Pacific Palisades, California: Goodyear.

Table 1. Descriptive and Correlation Results for Measures of Kinesthetic Sensitivity and Pursuit Rotor Scores Using the Nondominant (Left) Hand.

	Pursuit Rotor Descriptive Data		Kinesthetic Sensitivity Correlations with TOT Scores		
Blocks	Means	S.D.	JND	DEG	CM
1	3.95	2.80	-0.077	0.214	-0.087
2	5.79	3.12	0.021	0.191	-0.193
3	7.84	3.73	-0.018	0.191	-0.130
4	9.32	4.20	-0.185	0.002	-0.127
5	10.42	4.42	-0.192	0.037	-0.107
6	12.93	4.81	-0.107	0.138	-0.129
7	12.88	5.24	-0.105	0.115	-0.237
8	13.15	4.88	-0.166	0.144	-0.101
9	13.59	5.24	-0.151	0.154	-0.089
10	13.45	5.67	-0.170	0.108	-0.087
11	15.76	5.23	-0.162	0.130	-0.064
12	15.52	5.29	-0.130	0.071	-0.143
13	15.80	5.44	-0.182	0.060	-0.008
14	15.57	5.90	-0.324*	0.154	-0.018
15	15.83	6.00	-0.317*	0.043	-0.094

* Significant at .05 level

Table 2. Descriptive and Correlation Results for Measures of Kinesthetic Sensitivity and Pursuit Rotor Scores Using the Dominant (Right) Hand.

	Pursuit Rotor Descriptive Data		Kinesthetic Sensitivity Correlations with TOT Scores		
Blocks	Means	S.D.	JND	DEG	CM
1	9.45	5.44	0.027	-0.085	-0.077
2	13.53	5.83	0.033	-0.078	-0.063
3	15.83	5.84	0.052	-0.051	-0.093
4	16.23	6.57	0.051	-0.011	0.006
5	17.33	6.72	0.142	0.036	-0.059
6	18.92	7.07	0.190	-0.038	-0.017
7	19.60	7.10	0.155	0.012	-0.020
8	18.67	7.46	0.123	-0.061	-0.071
9	18.55	7.25	0.146	-0.044	-0.027
10	19.89	6.93	0.018	0.027	0.061
11	20.63	7.86	0.186	-0.083	0.044
12	20.37	7.21	0.186	-0.062	-0.013
13	20.77	6.71	0.121	-0.135	0.033
14	20.57	6.81	0.138	-0.073	0.048
15	20.94	7.03	0.098	0.033	0.025

Table 3. Descriptive and Correlation Results for Measures of Kinesthetic Sensitivity and Dart Tossing Scores Using the Nondominant (Left) Hand.

	Dart Tossing Descriptive Data		Kinesthetic Sensitivity Correlations with Error Scores		
Blocks	Means	S.D.	JND	DEG	CM
1	97.35mm	28.26mm	0.078	-0.075	0.446*
2	87.32	25.99	0.087	-0.187	0.074
3	83.79	24.25	-0.141	-0.144	0.185
4	82.52	25.10	-0.005	-0.032	0.040
5	87.01	27.47	0.011	-0.142	0.073
6	80.54	26.55	-0.123	-0.061	0.112
7	84.49	26.84	-0.178	0.040	0.188
8	82.51	27.54	-0.033	0.056	0.062
9	79.17	26.12	-0.200	-0.120	0.037
10	80.35	25.93	0.031	0.015	0.066

* Significant at the .05 level.

Table 4. Descriptive and Correlation Results for Measures of Kinesthet Sensitivity and Dart Tossing Scores Using the Dominant (Right) Hand.

	Dart Tossing Descriptive Data		Kinesthetic Sensitivity Correlations with Error Scores		
Blocks	Means	S.D.	JND	DEB	CM
1	67.82mm	23.86mm	-0.092	0.166	0.110
2	61.57	24.13	-0.247	0.117	0.114
3	57.22	25.04	-0.106	0.303*	0.171
4	57.60	21.26	-0.192	0.302*	-0.046
5	57.62	21.11	-0.162	0.352*	0.014
6	59.26	24.19	-0.014	0.397*	0.123
7	55.01	23.21	0.034	0.102	0.047
8	55.72	20.89	0.060	0.249	0.112
9	57.03	19.40	-0.098	0.271	0.043
10	53.38	19.01	-0.174	0.000	-0.021

* Significant at the .05 level.

FINAL REPORT NUMBER 17
RECEIVED A NO-COST TIME EXTENTSION
TO BE SUBMITTED IN 1987 MINI-GRANT FINAL REPORT
Ms. Jennifer Davidson (GSSSP)
760-6MG-024

FINAL REPORT NUMBER 18
REPORT NOT RECEIVED IN TIME
WILL BE PROVIDED WHEN AVAILABLE
Dr. Vito DelVecchio
760-6MG-076

INVESTIGATION OF PULSED DISCHARGES IN NITROGEN FOR PLASMA PROCESSING

A Final Scientific Report

Submitted by:

S. K. Dhali, Principal Investigator

December 31, 1987

Covering the Contract Period:

January 2, 1987 through December 31, 1987

Contract No. F49620-85-C-0013/SB5851-0360

Department of Electrical Engineering
Southern Illinois University
Carbondale, IL 62901

ABSTRACT

This report presents the results of computer simulations of pulsed discharges in nitrogen gas. A model which includes the external circuit, discharge tube dimensions, vibrational master equation, and the rate equations for the excited nitrogen molecules and nitrogen atoms has been developed. A Monte Carlo code for obtaining the electron distribution function, and an Adam's predictor-corrector code for the solution of the rate equations have been written. The concentration of discharge species under different discharge conditions have been obtained.

Investigation of Pulsed Discharges in Nitrogen for Plasma Processing

S. K. Dhali
Department of Electrical Engineering
Southern Illinois University
Carbondale, IL 62901

The goal of this research project, as stated in the proposal work statement were to A) model a pulsed discharge in nitrogen including the afterglow region; B) develop a Monte Carlo and Adam's predictor-corrector code; C) interpret the results in relation to the physical characteristics of the discharge. All goals were accomplished.

This study has provided fundamental knowledge regarding the rate of populating the different vibrational states and electronically excited states of nitrogen molecules, which has its importance in characterizing gas lasers, high power switches and plasma processing devices.

A part of this study was presented at the 40th Annual Gaseous Electronic Conference in Atlanta.

I. ACCOMPLISHMENT

The model developed is for pulsed discharges in nitrogen gas at pressures in the range of 1 - 100 Torr, reduced electric field (E/N) in the range of 100 - 200 Td, and discharge currents in the range 0 - 5 A. The model is valid for times well into the ms range.

A. DISCHARGE MODELING

Under the conditions specified, the discharges are weakly ionized which are sustained by an external electric field. The applied field, coupled with the relatively low rate of energy transfer between electrons and the much more massive neutral species, results in highly energetic electrons (10^4 K) and room temperature neutral species. Ion's transfer energy readily in elastic collisions with neutral species and are therefore usually near the neutral species temperature. Because of the difference in the electron and the neutral species temperature, the discharge is not at local thermodynamic equilibrium [1].

1. Circuit

The physical configuration of the discharge tube is shown in Figure 1. A step voltage (V_o) is applied to the electrodes in series with a resistor R. In our model we neglect all electrode effects. This is a reasonable assumption for short duration discharges where there is no sheath formation.

In our model the discharge region is treated as uniform in the axial direction. This is a resonable assumption if we assume the whole region to be preionized by the same amount. This simplifies the computation due to the absence of any spatial variations. Unless the gap is preionized, the discharge is usually initiated by a single electron. Computationally to follow the evolution of a discharge from a single electron becomes intractable. In all our

simulations we assume a preionization of the whole volume in the range of 10^6-10^9 cm^{-3}.

The radial profile of the electron distribution is assumed to be the fundamental mode of diffusion

$$J_o(2.405\ r/R_d),$$

where J_o is the first order Bessel's function, r is the radial distance and R_d is the radius of the discharge tube [2]. This is an approximation which is valid only when ambipolar diffusion sets in. The validity of this assumption is not critical at early times when the electron density is low and the loss due to diffusion is negligible. In the ms time scale the electron and ion densities are high enough for ambipolar diffusion to determine the radial density profile.

Since electron excitation is the primary cause for the production of other species, to a first approximation we assume the radial profile of all excited species to be the fundamental mode of diffusion. By integrating in the radial direction and using average densities, we have reduced the problem to zero dimension in space. The diffusion equation for electrons and excited species reduce to

$$dn_e/dt = (5.76\ n_e\ D_a)/\ R_d^2$$

and

$$d(N_2^*)/dt = (5.76\ N_2^*\ D_n)/\ R_d^2$$

respectively, where n_e and N_2^* are the average electron and neutral excited species density respectively, D_a is the ambipolar diffusion

coefficient, and D_n is the diffusion coefficient for molecular nitrogen.

For ion and neutral gas temperature of 500 K and gas pressure P in Torr, we used [3]

$$D_a = 2823/P,$$

and

$$D_n = 223/P.$$

The gap voltage V_g and the gap current I_g are determined from

$$dV_g/dt = (V_o - V_g - I_g R)/CR,$$

and

$$I_g = 1.36\, q\, R_d^2\, n_e V_g/d$$

respectively, where C is the gap capacitance, and d is the gap distance.

The gap capacitance, C, was assumed to be that due to a parallel-plate gap of length d, and electrode radius of 1 cm.

2. Discharge Processes

In order to obtain the time evolution of the discharge species, the electron energy distribution function (EEDF), the vibrational population (VP) and the density of the electronically excited species have been calculated self-consistently under given circuit parameters. The coupling between the EEDF and the VP, due to stepwise excitation or superelastic collisions have been taken into account since it can strongly modify the energy balance.

The method of calculation being followed is similar to the one being used by Boeuf and Kunhardt [4]. Since the relaxation times for electron-molecule and molecule-molecule reactions are very different, it is not necessary to calculate the electron energy distribution for the electrons and the rate equation for the excited species with the same time step. The time evolution of the electron distribution function is obtained by assuming successive quasi-equilibrium. That is, the electron distribution function is calculated at times t_0, t_1,...t_i,... assuming a state of equilibrium with the density of excited species at each of these times. During the time interval (t_i, t_{i+1}), the system rate equation is solved with the electron source term kept constant and equal to its value at time t_i. The time intervals are chosen small enough to allow only small changes in the electron distribution function. The time necessary for the electrons to reach equilibrium at a given time t_i is negligible with respect to the time intervals (t_i, t_{i+1}).

Described below are the reactions considered when formulating the rate equation for the vibrational population and the excited electronic state densities.

The vibrational excitation and de-excitation (superelastic collision) by electron impact is given by [5]

$$e + N_2(X^1\Sigma_g^+,v) \rightleftharpoons N_2(X^1\Sigma_g^+,\omega) + e$$

where $0 \leq v \leq 8$ and $0 \leq \omega \leq 8$, e is an electron, and $N_2(X^1\Sigma_g^+,v)$ [$N_2(v)$] represents a nitrogen molecule vibrationally excited in the

$$N_2(C) \rightarrow N_2(B) + h\nu$$

$$N_2(B) \rightarrow N_2(A) + h\nu$$

$$N_2(B) + N_2(X) \rightarrow N_2(A) + N_2(X)$$

$$N_2(B) + N_2(A) \rightarrow N_2(C) + N_2(X)$$

$$N_2(A) + N_2(A) \rightarrow N_2(B) + N_2(X) \text{ (Self Quenching)}$$

$$N_2(A) + N_2(A) \rightarrow N_2(C) + N_2(X) \text{ (Self Quenching)}$$

$$N_2(A) + N(^4S) \rightarrow N_2(X) + N(^2P)$$

$$N(^4S) + N(^4S) + N_2(X) \rightarrow 3N(X)$$

where $N(^2P)$ is one of the excited states of nitrogen atom and $h\nu$ is the photon energy.

The time evolution of the vibrational population and electronic states densities is described by the following system of master equations, where N_ν is the concentration of $N_2(X^1\Sigma_g^+,\nu)$:

$$(dN_\nu/dt) = (dN_\nu/dt)e\text{-}V + (dN_\nu/dt)V\text{-}V + (dN_\nu/dt)V\text{-}T + (dN_\nu/dt)A\text{-}V - 1283\ N_\nu/P \qquad (1)$$

The electron source term $(dN_\nu/dt)e\text{-}v$ is given by:

$$(dN_\nu/dt)e\text{-}v = n_e \sum_{w=0}^{8} (N_\nu K_{\omega\nu} - N_\nu K_{\nu\omega})$$

where $K_{\omega\nu}$ is the rate coefficient obtained from Monte Carlo simulation.

The second term in the right-hand side (rhs) of Eq. 1 is the rate of change of the N_ν due to V-V (vibrational - vibrational) reactions:

$$(dN_\nu/dt)V\text{-}V = \sum_{\omega=1}^{\nu^*+1} (N_{\nu+1}N_{\omega-1}P^{\omega-1,\omega}_{\nu+1,\nu} - N_\nu N_\omega P^{\omega,\omega-1}_{\nu,\nu+1})$$

$$+ \sum_{\omega=0}^{\nu^*} (N_{\nu-1}N_{\omega-1}P^{\omega+1,\omega}_{\nu-1,\nu} - N_\nu N_\omega P^{\omega,\omega-1}_{\nu,\nu-1})$$

The third term in the rhs of Eq. 1 is the rate of change of N_ν due to V-T (vibrational and translational) reactions:

$$(dN_\nu/dt)_{V-T} = N_2(N_{\nu+1}P_{\nu+1,\nu} - N_\nu P_{\nu,\nu+1}) + N_2(N_{\nu-1}P_{\nu-1,\nu} - N_\nu P_{\nu,\nu-1})$$

The rate coefficients for the V-V and V-T reactions $P^{\omega,\omega-1}_{\nu,\nu+1}$ and $P_{\nu,\nu-1}$ respectively are obtained from Bray [6], Capitelli and Dilonardo [7].

The fourth term in the rhs of Eq. 1 is the rate of change of N_ν due the self-quenching of the A-state. It is given by:

$$(dN_\nu/dt)A\text{-}V = N_A^2(K^{XB}_{AA}\delta_{\nu,8} + K^{XB}_{AA}\delta_{\nu,2})$$

where $\delta_{i,j} = 1$, if $i=j$ and $\delta_{i,j} = 0$, if $i \neq j$

The last term in the rhs of Eq. 1 is the loss of N_ν due to diffusion process.

Equation 1 has been coupled with rate equations for the A (N_A), B (N_B) and C (N_C) states, electron (n_e), nitrogen atom (N) and discharge voltage (V_g) to form a set of 51 coupled ordinary differential equations.

$$(dN_A/dt) = n_e N_2 K_{eA} + N_2 N_B K^{AX}_{BX} + N_B K^A_B - 2N_A^2(K^{BX}_{AA} + K^{CX}_{AA}) - (N_A N_B K^{CX}_{AB}) - N_2 N_A K^{BX}_{AX} - 1283\, N_A/P$$

$$(dN_B/dt) = n_e N_2 K_{eB} + N_C K^B_C - N_B K^A_B - N_2 N_B K^{AX}_{BX} - N_A N_B K^{CX}_{AB} + N_A^2 K^{BX}_{AA} - 1283\, N_B/P$$

νth level of the ground state. Higher vibrational level reactions are assumed to play a minor role.

Electronic excitation by electron impact is assumed to occur only from the ground state $N_2(X\,^1\Sigma_g^+,\ \nu=0)$. The electronic excitation processes considered are,

$$N_2(X) + e \rightarrow N_2(A) + e$$

$$N_2(X) + e \rightarrow N_2(B) + e$$

$$N_2(X) + e \rightarrow N_2(C) + e$$

$$N_2(X) + e \rightarrow N_2^+ + 2e$$

$$N_2(X) + e \rightarrow N(^4S) + N(^4S)$$

where $N_2(A)$, $N_2(B)$ and $N_2(C)$ are the electronically excited states of N_2, and $N(^4S)$ is the ground state nitrogen atom.

Several types of molecule-molecule collisions have been taken into account. However, only the collisions between vibrationally excited molecules in the ground state are considered. In addition, reactions involving change in more than one vibrational quantum is neglected. The following reactions between vibrationally excited molecules are considered:

$$N_2(\nu) + N_2(\omega) \rightarrow N_2(\nu+1) + N_2(\omega-1),$$

where $0 \leq \nu \leq \nu^*-1$, $0 \leq \omega \leq \nu^*$ [$\nu^* = 45$ being the highest vibrational level and (ν^*+1) correspond to the dissociated molecule, and

$$N_2(\nu) + N_2(\omega) \rightarrow N_2(\nu-1) + N_2(\omega).$$

The following is a list of reactions, involving the electronically excited states, that are included in the model:

$$(dN_C/dt) = n_e N_2 K_C + N_A^2 K_{AA}^{CX} - N_C K_C^B + N_A N_B K_{AB}^{CX} - 1283\, N_C/P$$

$$(dn_e/dt) = n_e N_2 q_i - (16261/P)\,(n_e) - 2K_{e+} n_e^2$$

$$(dN/dt) = (dN_{v=47}/dt) + n_e N_2 K_{eN} - 2N^2 N_2 K_N$$

$$(dV_g/dt) = (1/RC)\,[V_o - V_g - (1262\times10^{-18} \times 4\times10^3 N_e V_g)]$$

where q_i, K_{eA}, K_{eB}, K_{eC}, and K_{eN} are obtained from the Monte Carlo simulation. The rate coefficients $K_{AA}^{CX} = 1.1\times10^{-9}$ cm^3/s and $K_{AA}^{BX} = 2.6\times10^{-10}$ cm^3/s are taken from the results of Hays and Oskam [8]. The rate coefficients $K_{AX}^{BX} = 1.8\times10^{-12}$ cm^3/s, $K_{BX}^{AX} = 1.3\times10^{-11}$ cm^3/s, $K_{AB}^{CX} = 4.6\times10^{-10}$ cm^3/s, $K_B^A = 2.0\times10^5$ s^{-1}, and $K_C^B = 2.2\times10^7$ s^{-1} are taken from the paper by Pivovar and Sidorova [9]. We assume that only $N(^4S)$ atoms can recombine to form N_2 via the three body process, and the value of the rate coefficient $K_N = 8.3\times10^{-34}$ exp (500/T) for this reaction, where T = 500 K [10]. The global electron recombination rate is given by the following functional forms [10]:

for $1\times10^{-17} \le E/N \le 4\times10^{-17}$ V-cm^2:

$$K_{e+} = \exp\,(-13.28 - (7.32\times10^{16}\ E/N))$$

for $4\times10^{-17} \le E/N \le 2.2\times10^{-16}$ V-cm^2:

$$K_{e+} = \exp\,(-16.15 - (1.32\times10^{15}\ E/N)).$$

Except for the electron impact dissociation cross section, which was obtained from the experimental results of Zipf and McLaughlin [11], all other electron impact cross sections used were taken from Phelps and Pitchford [12].

2. CODE DEVELOPMENT

We have developed a Monte Carlo code for obtaining the electron impact rate coefficients. The code takes into account the elastic, vibrational, electronic, dissociating, ionizing, and superelastic collisions. The code has been verified against measured values of drift velocity and rate coefficients under equilibrium conditions.

The Adam's predictor-corrector method implemented is a modification of the GEAR routine which is capable of handling stiff problems [13].

The routines, in the form of FORTRAN codes, are available upon request.

C. RESULTS AND DISCUSSIONS

We present here the results of simulations using the model developed in the previous section. The data presented is for a wide range of pressures, applied voltages, and preionization densities. We have looked at the time evolution of E/N in the bulk of the discharge, discharge current, electron density, and nitrogen A-state density. Although the time evolution of the other excited states is readily available from the simulation, in the first part of this section we have only concentrated on the A-state due to its practical importance.

At the end of this section we discuss some of the more recent calculations on a supercomputer in which the excited states are dealt with in more detail.

1. Superelastic Collisions

Shown in Figure 2 is the time evolution of the nitrogen A-state density for two cases: 1) superelastic collisions were not included in the simulation, 2) superelastic collisions were included in the simulation. The discharge corresponding to Figure 2 was for a pressure of 10.0 Torr, limiting discharge current of approximately 4.5 mA, series resistor of 4.0 kΩ, and step voltage of 25 kV.

The superelastic collisions start to influence the discharge at approximately 100 μs. This is approximately the time it takes for the vibrational population to reach a significant fraction of the total population as shown in Figure 3. The lower vibrational levels ($v \leq 8$) reaches a constant value well before the whole vibrational manifold has reached equilibrium. The population of the first few vibrational levels follow closely a Boltzmann distribution, the same is not true for the higher vibrational levels.

Figure 4 shows the effect of superelastic collisions on the EEDF. It is clear from the figure that there is a drastic increase in the tail of the distribution due to superelastic collisions. These results are in good agreement with those of Boeuf et al. [4] and Capitelli et al. [5].

It is apparent from these calculations that superelastic collisions have to be taken into account for the simulation results to be valid beyond a few microseconds. In all the subsequent results, superelastic collisions have been included for the evaluation of the EEDF.

2. Voltage Variations

Several calculations were done for voltages ranging from 20 to 30 kV, a pressure of 10 Torr, a series resistance of 4.0 kΩ, and a preionization density of 10^8 cm^{-3}. The results for 20, 25, and 30 kV are shown in Figure 5. As anticipated, the maximum E/N achieved during the discharge is higher for the higher applied voltage. Since the higher maximum E/N results in higher ionization rate, the build-up of the discharge current is faster for the higher applied voltage as shown in Figure 5. As a result of the faster buildup of the discharge current, the discharge voltage collapses to its steady state value much quicker for a higher applied voltage as compared to a lower voltage.

From the plots of the discharge voltage, it can be seen that for all three voltages the "steady-state E/N" is approximately the same (~ 50 Td). This is an expected result because the steady-state E/N is determined by the electron loss rate, which is same for all applied voltages. The loss of electrons in our model is due to diffusion to the walls. The oscillations near steady-state is due to the sensitivity of the ionization rate coefficient to the tail of the EEDF at low E/N. These fluctuations can be reduced by running the Monte Carlo for a larger set of electrons and by updating the rate coefficients more often. Any effort to reduce these oscillations increases the required computation time considerably. Since we are only interested in mean values at steady state, the improvements in the reduction of the fluctuations is only cosmetic.

In the more complete model, which is discussed later in this section, we have included the loss of electrons due to three body recombination.

The EEDF at 0.5 ms from initiation is shown in Figure 6 for the three different applied voltages. They are almost identical except for the small shift of the distribution towards the high energy end for higher applied voltages. The population of the vibrational manifold, which is shown in Figure 7, is partly responsible for the slight differences in the EEDF because of the superelastic collisions.

The plots shown in Figure 7 are at 5.0 ms from the initiation of the discharge and normalized with respect to the population of the ground-state nitrogen molecule. For the higher of the three voltages (30 kV) the vibrational manifold has almost reached equilibrium at 0.5 ms, whereas for 20 kV the vibrational population only extends up to the 20th quantum level. The time evolution of the vibrational manifold for an applied voltage of 25 kV is shown in Figure 8. By 1.0 ms the vibrational population up to the 23rd quantum number have reached a steady state, and by 5.0 ms all the vibrational levels have reached an equilibrium value. In our model, only the first eight vibrational levels are populated by electron impact, the rest of states are populated by V-V and V-T collisions. The shape and the time scale of evolution of the vibrational population are similar to that obtained by Boeuf et al. [4] and Capitelli et al. [5]. A direct comparison is not possible because

in our simulations the E/N was dictated by the external circuit, whereas in literature most of the work so far has been done for a constant E/N.

Figure 9 shows the time evolution of the $N_2(A)$ for the three different applied voltages. The A-state population is directly related to evolution of the E/N and growth of the discharge current. Due to a higher E/N and a faster growth of discharge current at 30 kV, the A-state density reaches a steady state earlier as compared to a lower applied voltage. The final steady-state value of the A-state density is approximately 5.0×10^{13} cm^{-3} for all three voltages. In the simulations presented in this subsection the quenching of the A-state due to nitrogen atoms was neglected. Moreover, it appears from the simulations done by Capitelli et al. [5] that nitrogen atoms reach significant proportions only after about 100 ms.

3. Pressure Variations

We have done a set of simulations with pressures ranging from 1-100 Torr. To make any meaningful comparison for different pressures, the applied voltage and the series resistor were scaled to give approximately the same current at steady state. For pressures of 1, 10, and 100 Torr, the applied voltage and series resistor used were 2.5 kV and 0.4 kΩ, 25kV and 4.0 kΩ, and 250 kV and 40.0 kΩ respectively.

Figure 10 shows the time evolution of E/N for the three different pressures. The rise time difference in the three plots is due to scaling of the resistors. Although the applied voltages are

different, the maximum E/N achieved in each case is about the same due to scaling of the pressure. The E/N required to sustain the discharge is higher for lower pressure. This is primarily due to increased diffusion loses at low pressures.

Figure 11 shows the time evolution of discharge current for the three different pressures. The currents are in the range of 3 - 4 mA and towards the low range for lower pressures. The lower currents for lower pressures is again due to the increased loss of electrons due to diffusion to the wall.

The normalized vibrational population at 0.5 ms from initiation of the discharge is shown in Figure 12. Due to low V-V and V-T pumping rates at low pressures, the vibrational population is far from its equilibrium values. Partly for the same reason and partly because of higher sustaining E/N at low pressures, the fractional population of the lower vibrational quantum numbers is higher as compared to a higher pressure.

The EEDF for the three pressures is shown in Figure 13 for 0.5 ms from initiation. Because of a higher E/N for low pressures, the tail of the distribution extends more towards the high energy end. The tail of the EEDF is also enhanced by the higher number of superelastic collisions due to increased fractional population of the first eight vibrational levels at low pressures.

4. Effect of Preionization

We have done a set of simulations with preionization ranging from 10^6 - 10^{10} cm^{-3}. For this set of simulation the pressure was

100 Torr, the applied voltage was 250 kV, and the series resistor was 40 kΩ.

The time evolution of the electron density for preionization of 10^6, 10^8, and 10^{10} is shown Figure 14. Past the microsecond range, the electron density in all three cases are very similar. This is an expected result because all other parameters of the discharge are the same. A higher level of preionization essentially pushes the discharge towards equilibrium at a faster rate. This can also be seen from the time evolution of E/N for the three different preionization levels in Figure 15.

The time evolution of the nitrogen A-state density for the three different pressures is shown in Figure 16. The evolution is similar to that of the electron density. The A-state density builds up at a higher rate for high preionization, but for all preionization levels the density eventually reaches a constant value of approximately 10^{13} cm^{-3}.

5. Excited-State Density:

For this section, the complete model for pulsed nitrogen discharge described in the previous section was used for the simulation.

The excited A-, B-, and C-states of the nitrogen molecule for a pressure of 10 Torr, voltage of 25 kV, and series resistance of 4 kΩ are shown in Figure 17. These states tend to reach equilibrium slightly before the bulk E/N. At the initial stage of the discharge, when the E/N is relatively high, the population of B- and

(1) Superelastic collisions enhance the EEDF tail, resulting in much higher rate coefficients, and higher equilibrium vibrational population and A-state density;

(2) Higher applied voltages produce higher equilibrium vibrational population and A-state density;

(3) Higher pressures required less time for the vibrational population to reach equilibrium and also has lower equilibrium E/N value. However, lower pressures achieved higher equilibrium vibrational population and A-state density;

(4) Higher initial electron density reduces the time for the discharge to reach equilibrium but produces identical equilibrium values;

(5) A typical discharge at a current of 3.5 A requires a sustaining E/N of 60 - 70 Td, produces electron density of 5.0×10^{11} - 1.0×10^{12} cm^{-3}, and A-state density in the range of 5.0×10^{13} - 1.0×10^{14} cm^{-3}.

The Monte Carlo simulation program developed can be easily modified to accommodate any number of electron-molecule interactions and gas mixtures which occur most frequently in plasma processing. The next step would be to add trace amounts of a reactant like silane (SiH_4), and study its effect on the discharge. If the reaction pathways and rates are known, one can predict the rate of formation of various products.

LIST OF REFERENCES

1. J. Loureiro, C. M. Ferreira, "Coupled Electron Energy and Vibrational Distribution Functions in Stationary N_2 Discharges," J. Phys. D: Appl. Phys., 19, 17 (1986).

2. J. D. Cobine, Gaseous Conductors, Theory and Engineering Applications, New York, Dover, 1958.

3. A. J. Davies, C. S. Davies, and C. J. Evans, "Computer Simulation of Rapidly Developing Gaseous Discharges," Proc. IEE, 118, 816 (1971).

4. J. P. Boeuf and E. E. Kunhardt, "Energy Balance in a Nonequilibrium Weakly Ionized Nitrogen Discharge," J. Appl. Phys., 60, 915 (1986).

5. M. Capitelli, M. Dilonardo, and C. Gorse, "Coupled Solutions of the Collisional Boltzman Equation for Electrons and the Heavy Particle Master Equation in Nitrogen," Chem. Phys., 56, 29 (1981).

6. K. N. C. Bray, "Vibrational Relaxation of Anharmonic Oscillator Molecules: Relaxation under Isothermal Conditions," J. Phys. B, 1, 705 (1968).

7. M. Capitelli and M. Dilonardo, "Nonequilibrium Vibrational Population of Diatomic Species in Electrical Discharges: Effects on the Dissociation rates," Chem. Phys., 24, 417 (1977).

8. G. N. Hays and H. J. Oskum, "Population of $N_2(B)$ and $N_2(A)$ During Nitrogen Afterglow," J. Chem. Phys., 50, 1507 (1973).

9. A. V. Pivovar and T. D. Sidorova, "Numerical Analysis of Stepwise Ionization in a Self-Sustaining Bulk Discharge in Nitrogen," Sov. Phys. Tech. Phys., 30, 308 (1985).

10. R. C. Slater, H. A. Hymen, D. Trainor, and A. Flusberg, "Research on Sources of Gas Phase Metastable Atoms and Molecules," AFWAL Report AFWAL-TR-82-2029, May 1982.

11. E. C. Zipf and R. W. Mclaughlin, "On the Dissociation of Nitrogen by Electron Impact and by E.U.V. Photo-absorption," Planet. Space Sci. 26, 449 (1978).

12. A. V. Phelps and L. C. Pitchford, "Anisotropic Scattering of Electrons by N_2 and its Effects on Electron Transport: Tabulation of Cross Section and Results," JILA Information Center Report No. 26, May 1985.

13. C. W. Gear, "The Automatic Integration of Ordinary Differential Equations," Comm. ACM, 14, 176 (1971).

C-state are higher than the A-state. Due to the eventual decay of these two states into the A-state and relatively lower E/N, the population of the A-state is about an order of magnitude higher than the B-state and about three orders of magnitude higher than the C-state.

The time evolution of the A-, B-, and C-state of the nitrogen molecule for a pressure of 1 Torr, voltage of 2.5 kV, and series resistance of 400 Ω are shown in Figure 18. The structure is very similar to the one described for 10 Torr. Although the population of nitrogen molecules is an order of magnitude lower at 1 Torr as compared to 10 Torr, the population of the excited states are comparable due to the relatively higher "steady-state" E/N at lower pressures.

II. CONCLUSIONS

Simulations of the time evolution of the electron energy distribution and the vibrational population in weakly ionized nitrogen plasma confirm the coupling between EEDF and VP due to superelastic collisions which leads to an important increase in the tail of the EEDF. Consequently, the macroscopic parameters (excitation, dissociation and ionization rates) in these conditions are drastically different from those obtained in cold gas approximation (zero vibrational temperature). The following conclusions are drawn from the above results:

FIGURE CAPTIONS

1. Schematic of the discharge tube and external circuit.

2. Time evolution of the nitrogen A-state: 1) without superelastic collisions in the simulation; 2) with superelastic collisions in the simulation. The applied voltage (V_o) is 25 kV, the pressure is 10 Torr, and the series resistance is 4.0 kΩ.

3. Normalized Vibrational Population at 0.5 ms into the discharge: 1) without superelastic collisions; 2) with superelastic collisions. All other conditions are as in Figure 2.

4. Electron Energy Distribution Function at 0.5 ms into the discharge: 1) without superelastic collisions; 2) with superelastic collisions. All other conditions are as in Figure 2.

5. Time evolution of E/N for three different voltages: 1) 20 kV; 2) 25 kV; 3) 30 kV. The pressure used is 10 Torr and the series resistance is 4.0 kΩ.

6. Electron Energy Distribution Function at 0.5 into the discharge: 1) 20 kV; 2) 25 kV; 3) 30 kV. All other conditions are as in Figure 5.

7. Vibrational Population at 0.5 ms into the discharge: 1) 20 kV; 2) 25 kV; 3) 30 kV. All other conditions are as in Figure 5.

8. Electron energy distribution function for different times for an applied voltage of 25 kV. All other conditions are as in Figure 5.

9. Time evolution of the nitrogen A-state density: 1) 20 kV; 2) 25 kV; 3) 30 kV. All other conditions are as in Figure 5.

10. Time evolution of E/N: 1) pressure of 100 Torr, applied voltage of 250 kV, and series resistance of 40.0 kΩ; 2) pressure of 10 Torr, applied voltage of 25 kV, and series resistance of 4.0 kΩ; 3) pressure of 1 Torr, applied voltage of 2.5 kV, and series resistance of 400 Ω.

11. Time evolution of the discharge current. All other conditions are as in Figure 10.

12. Normalized vibrational population at 0.5 ms into the discharge. All other conditions are as in Figure 10.

13. Electron energy distribution function at 0.5 ms into the discharge. All other conditions are as in Figure 10.

14 Time evolution of the electron density for three different levels of preionization: 1) 10^6 cm^{-3}; 2) 10^8 cm^{-3}; 3) 10^{10} cm^{-3}. The applied voltage is 25 kV, the pressure is 10 Torr, and the series resistance is 4.0 kΩ.

15. Time evolution of E/N for three different levels of preionization: 1) 10^6 cm^{-3}; 2) 10^8 cm^{-3}; 3) 10^{10} cm^{-3}. All other conditions are as in Figure 10.

16. Time evolution of the nitrogen A-state density for three different levels of preionization: 1) 10^6 cm^{-3}; 2) 10^8 cm^{-3}; 3) 10^{10} cm^{-3}. All other conditions are as in Figure 10.

17. Time evolution of A-, B-, and C-state density for a discharge pressure of 10 Torr, voltage of 25 kV, and series resistance of 4.0 kΩ.

18. Time evolution of A-, B-, and C-state density for a discharge pressure of 1 Torr, voltage of 2.5 kV, and series resistance of 400 Ω.

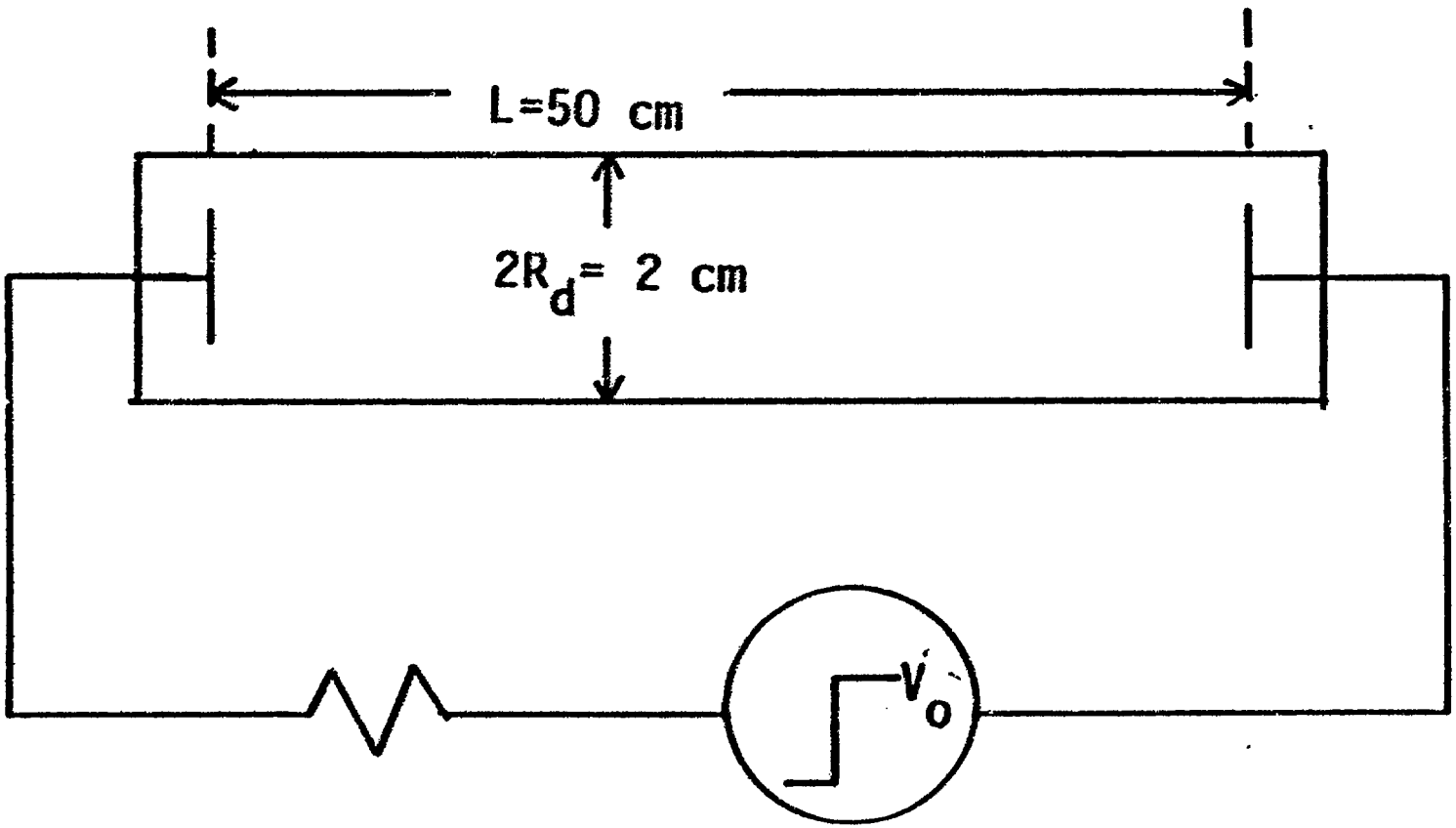
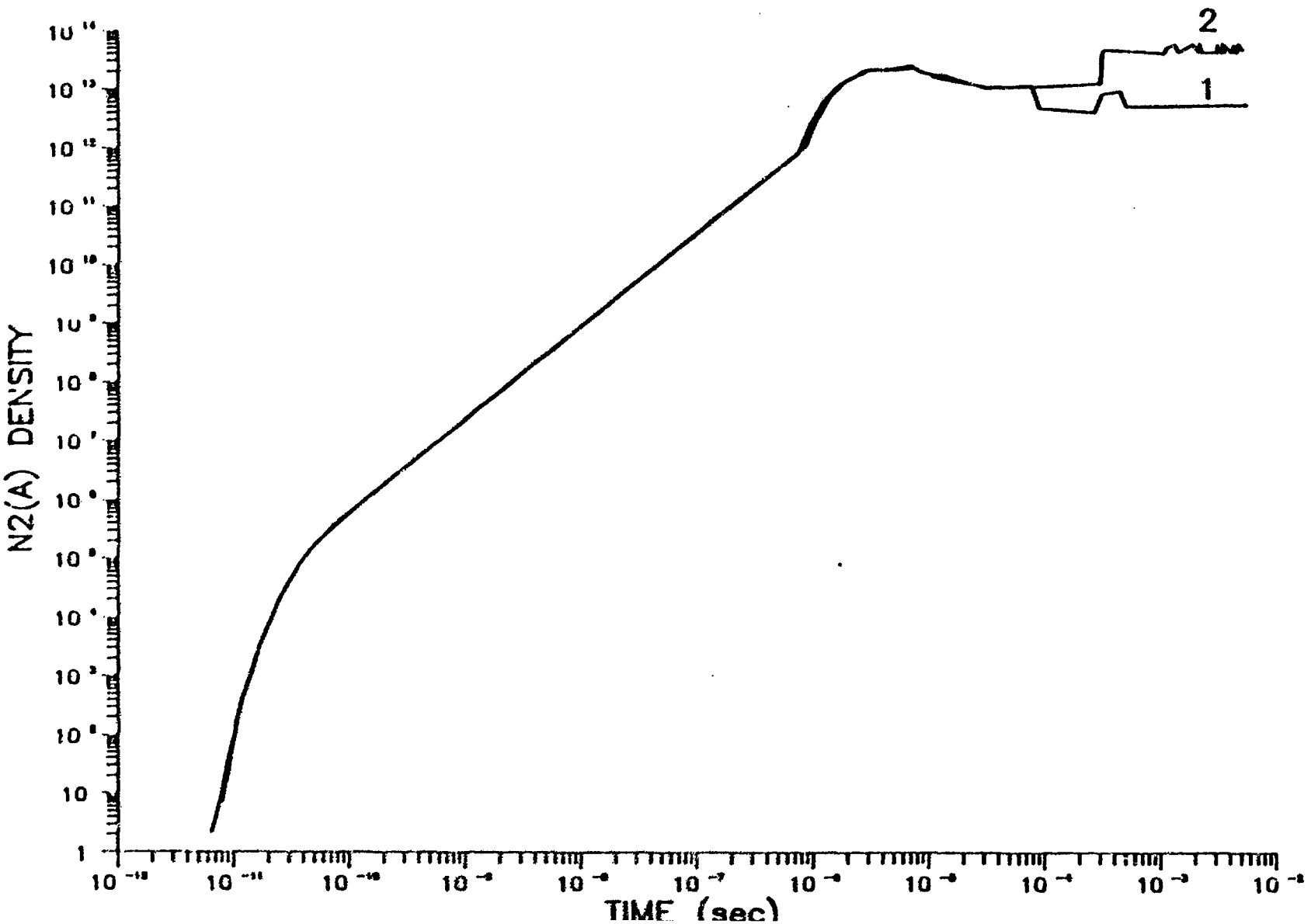
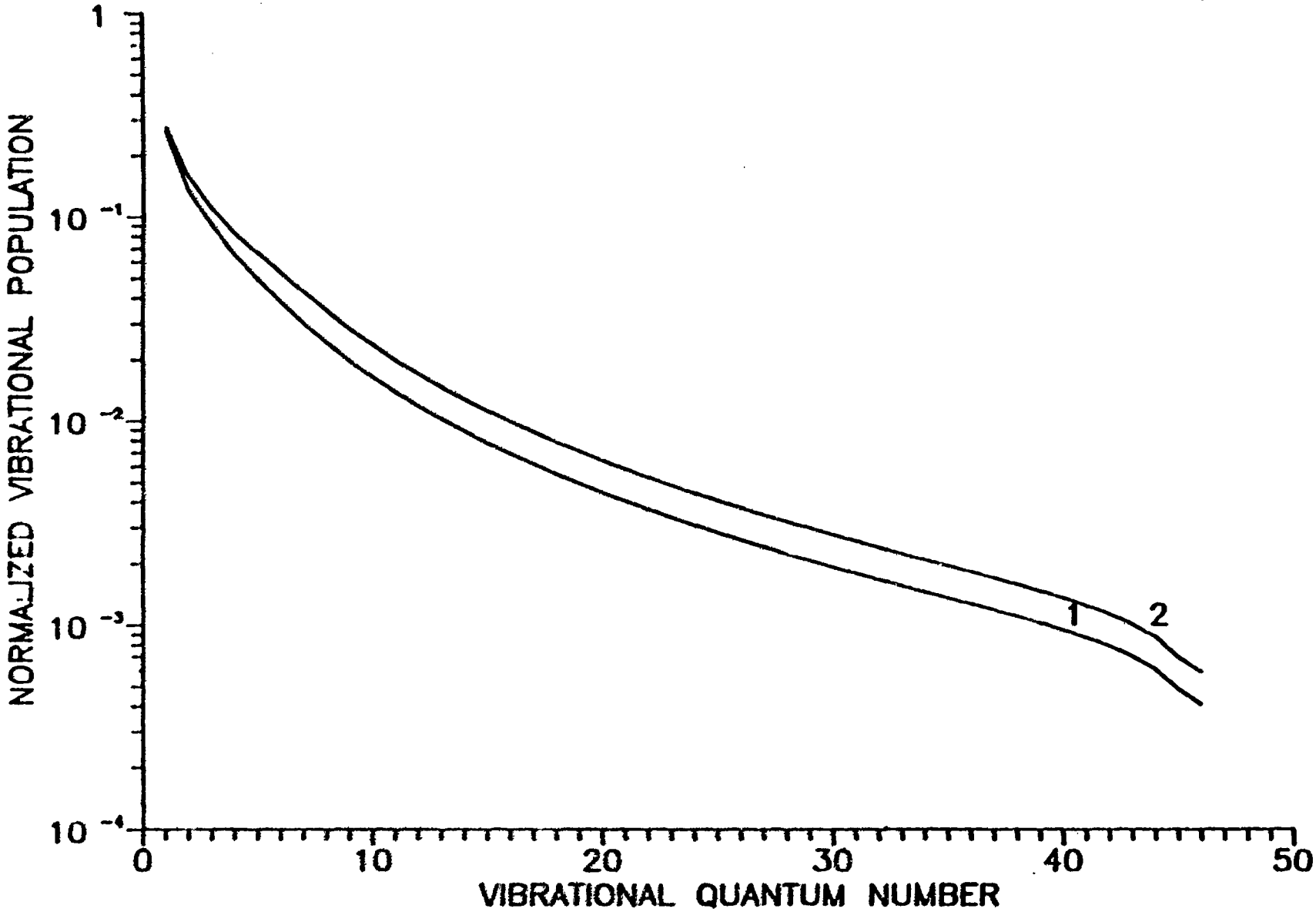


FIGURE 1

Time evolution of the nitrogen A-state density



Vibrational population at 0.5 msec



EDF at 0.5 msec

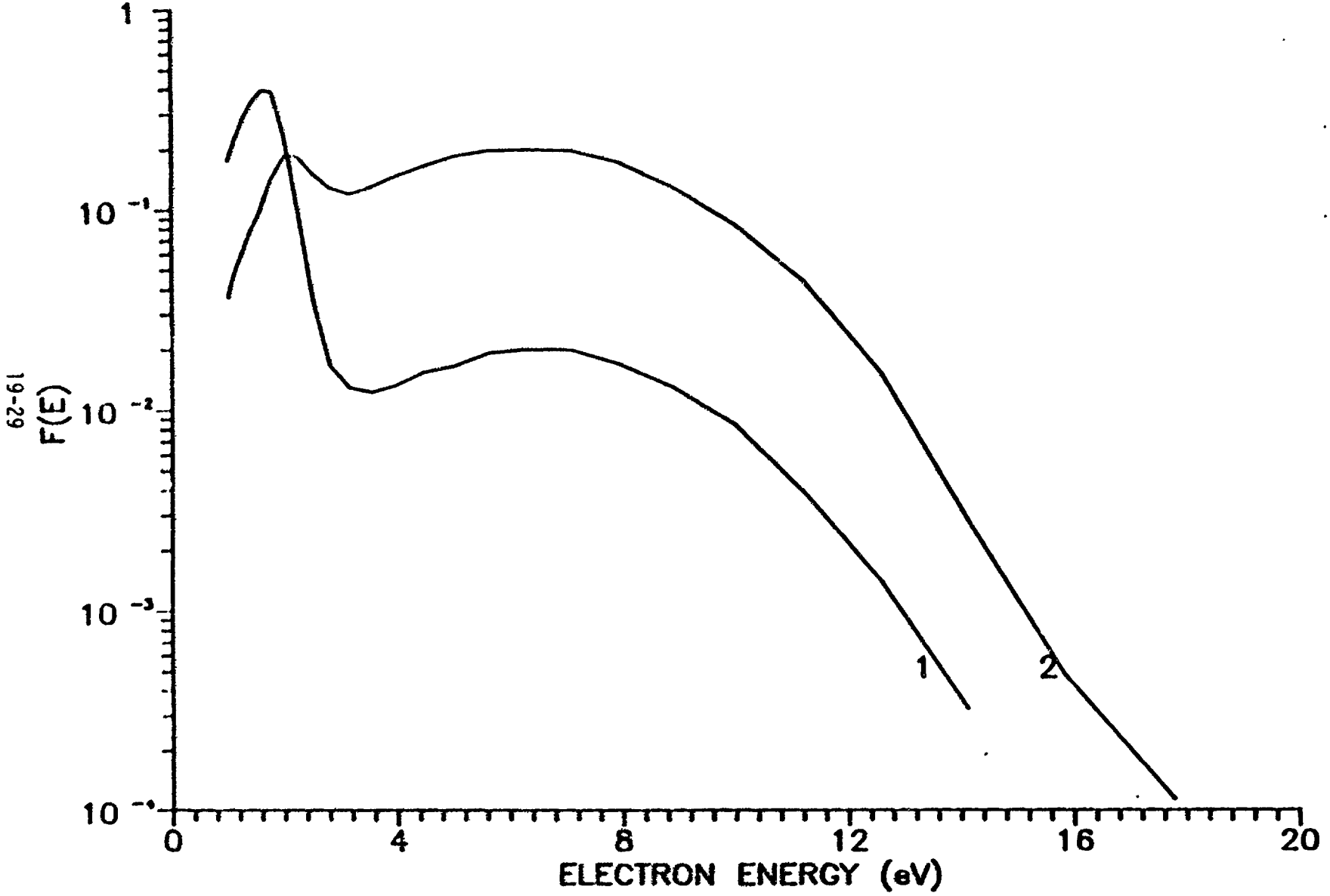
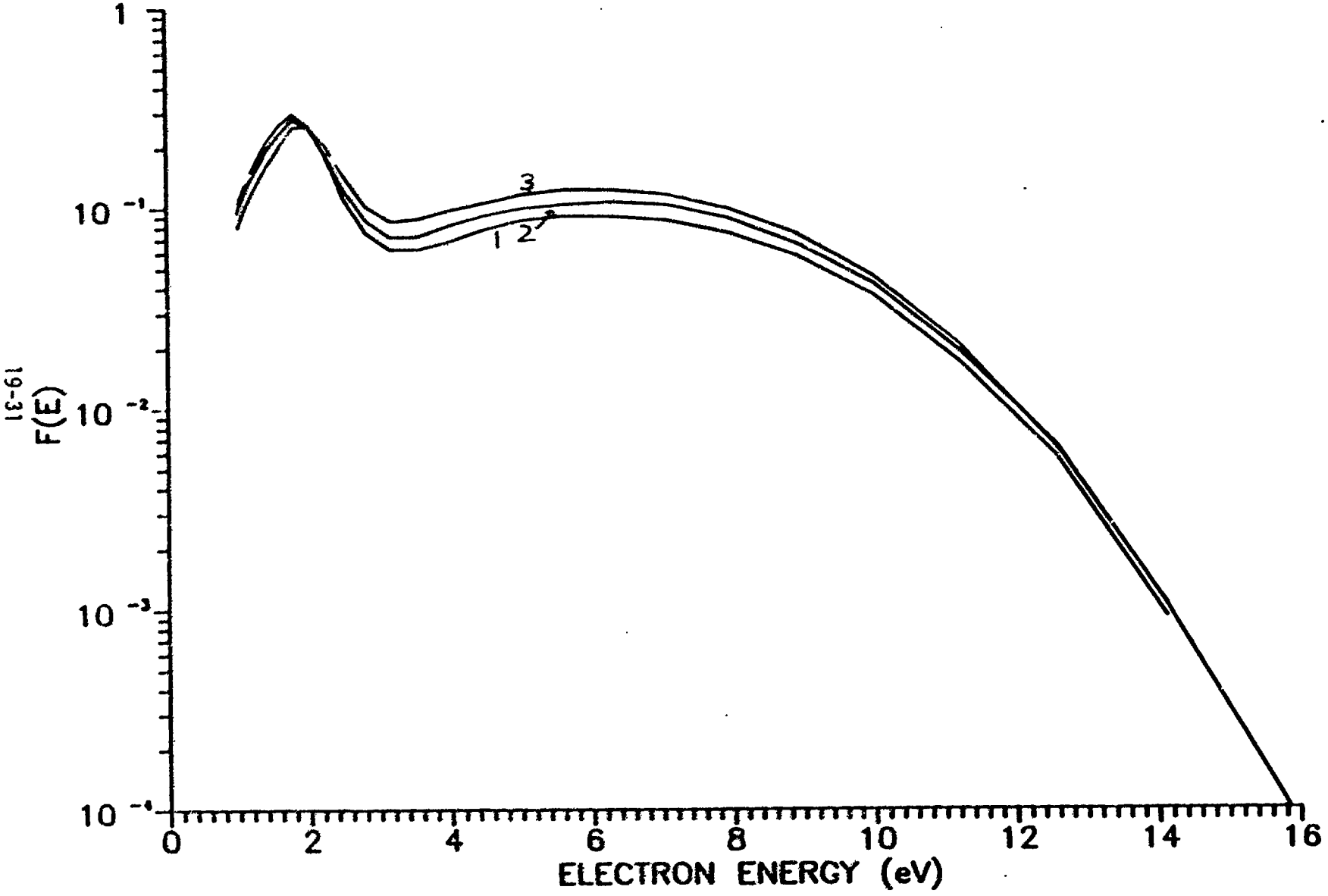


FIGURE 5

EDF at 0.5 msec



Vibrational population at 0.5 msec

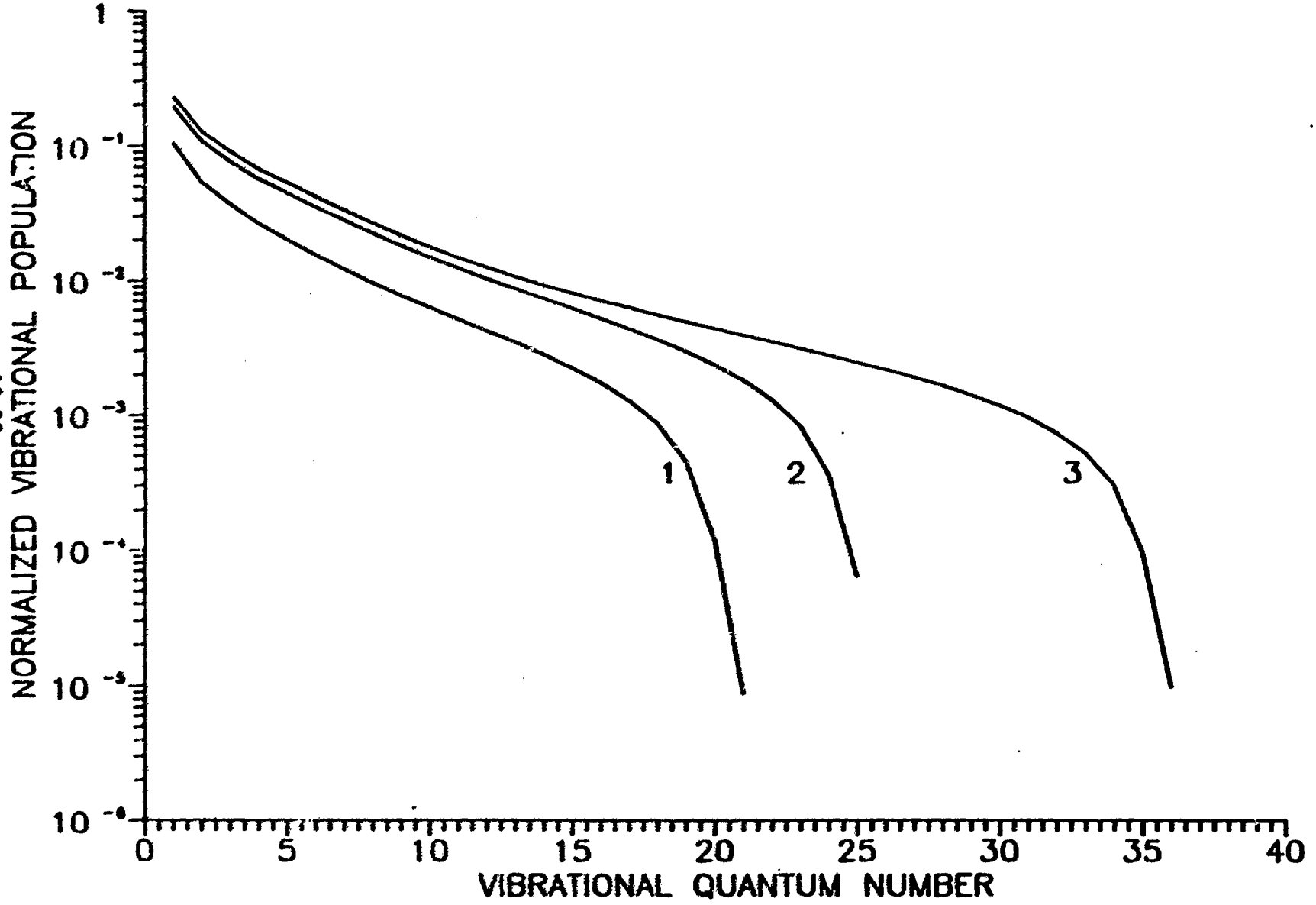


FIGURE 7

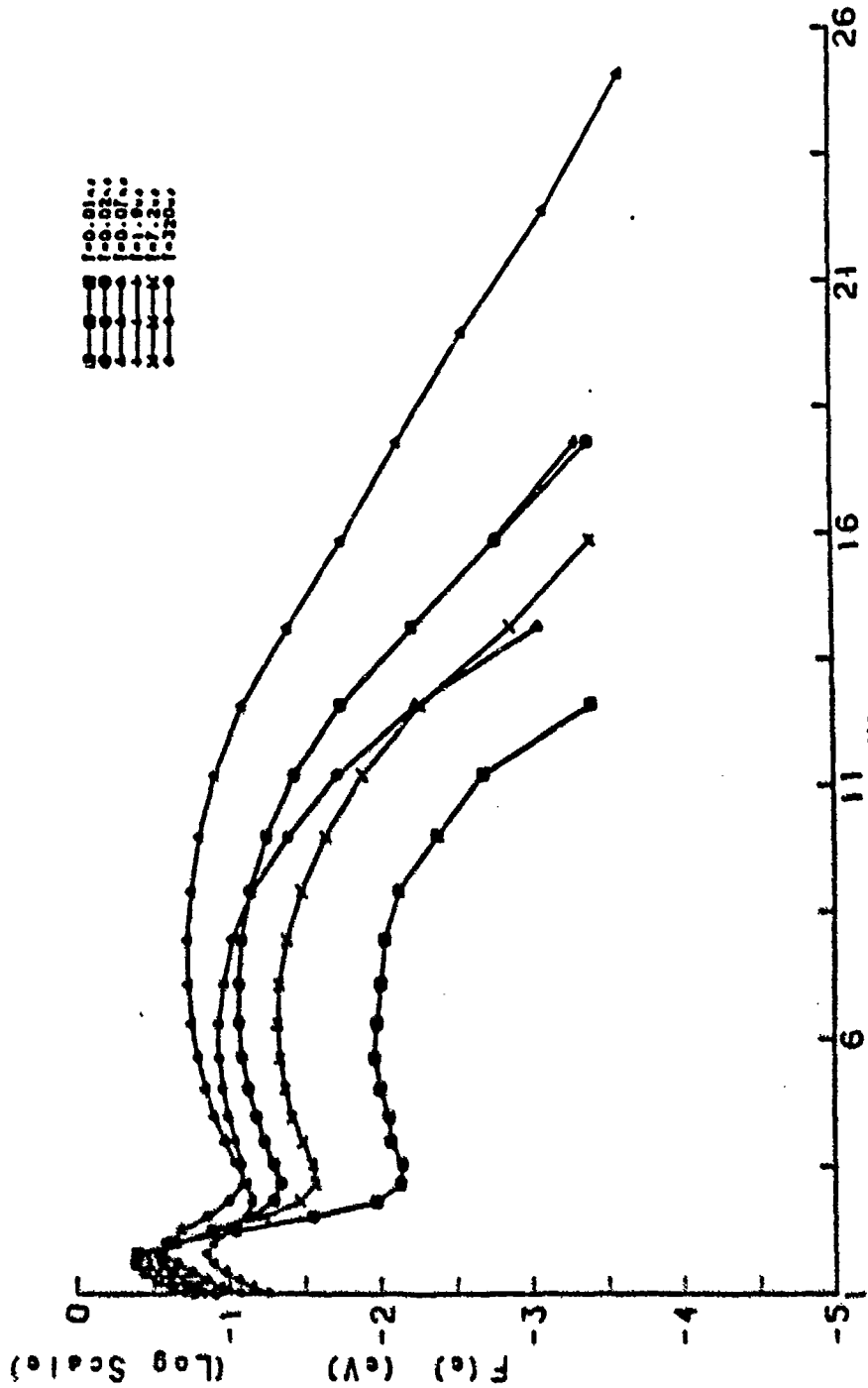


FIGURE 8

Time evolution of the nitrogen A-state density

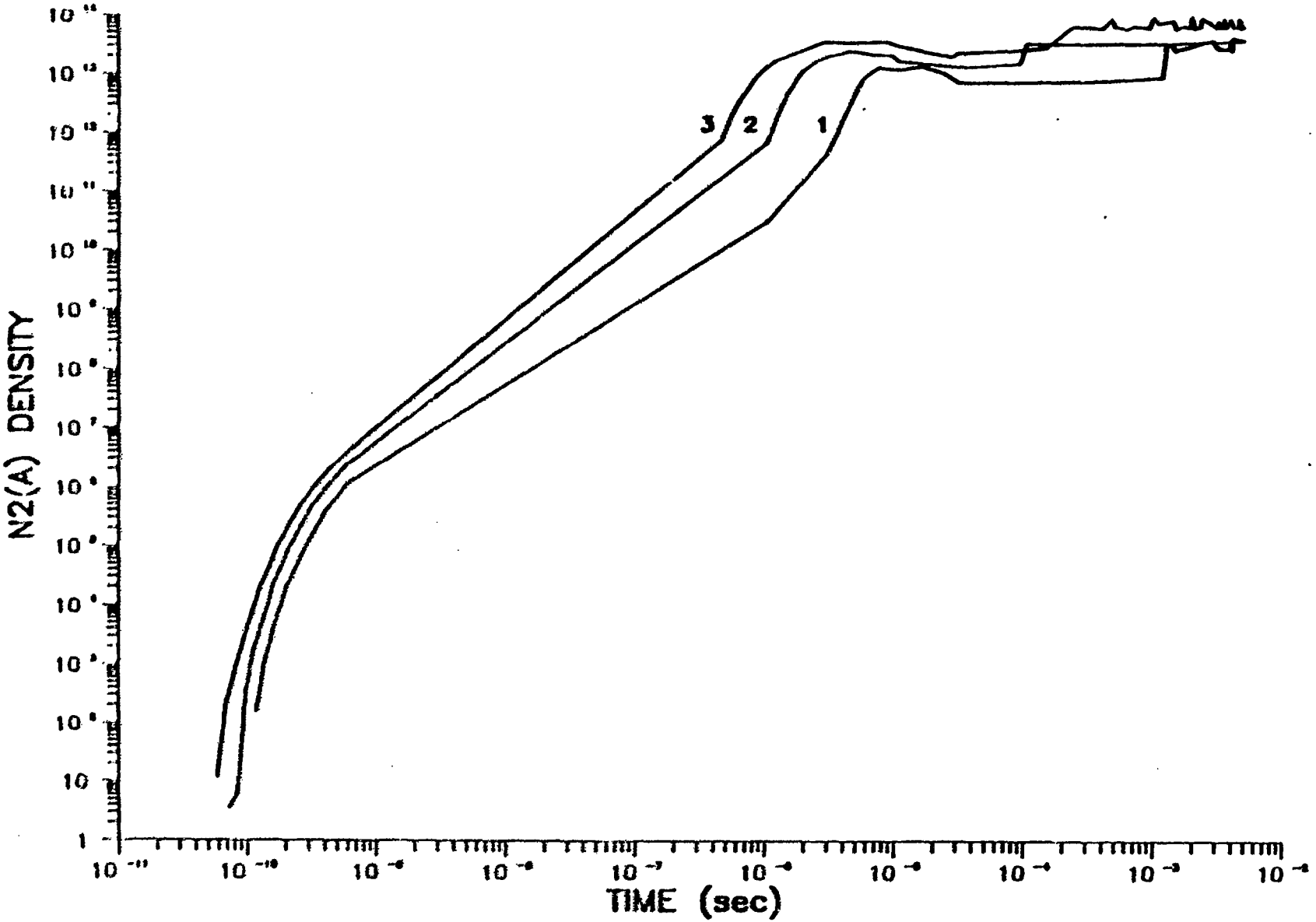
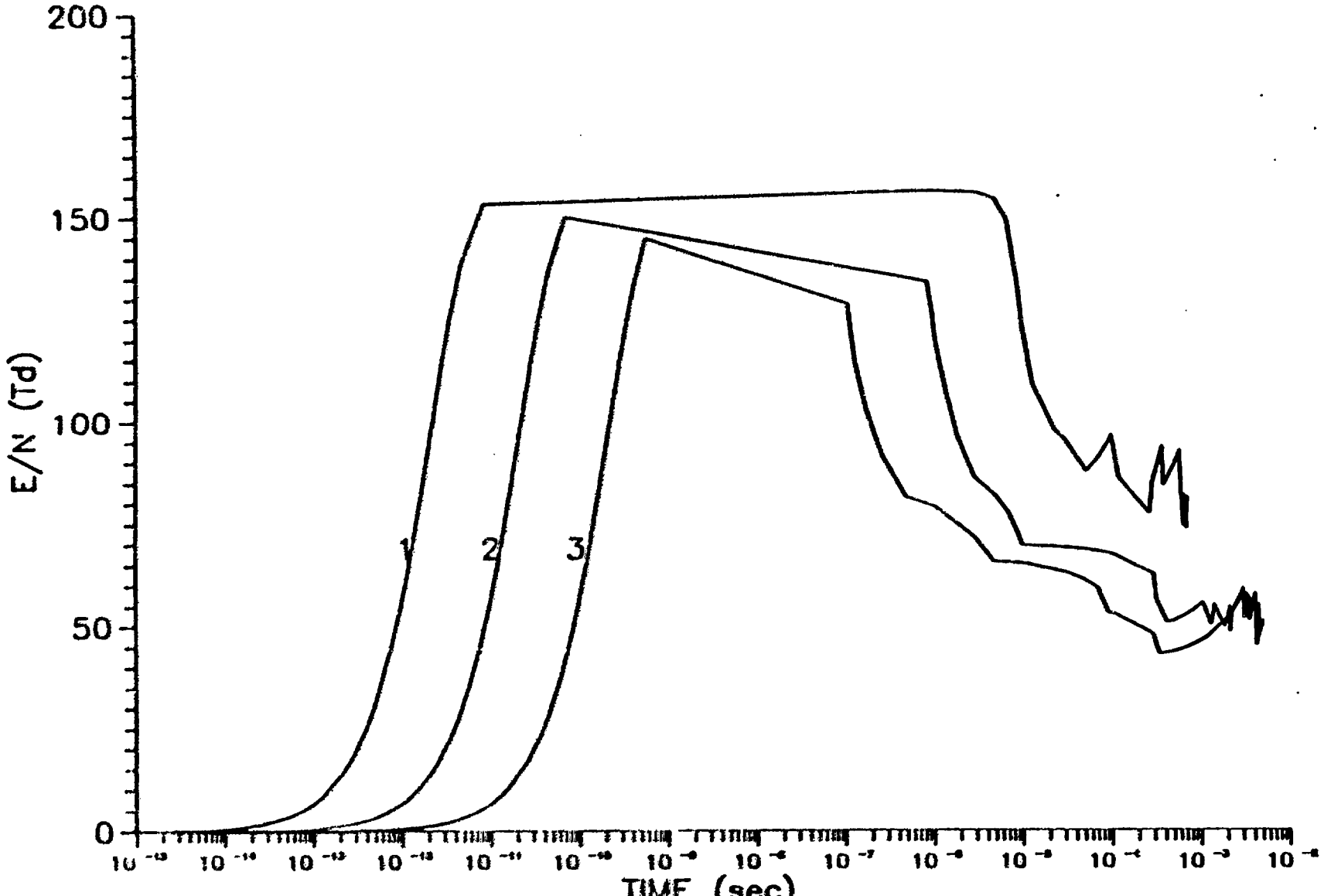


FIGURE 9

Time evolution of the E/N



Time evolution of the discharge current

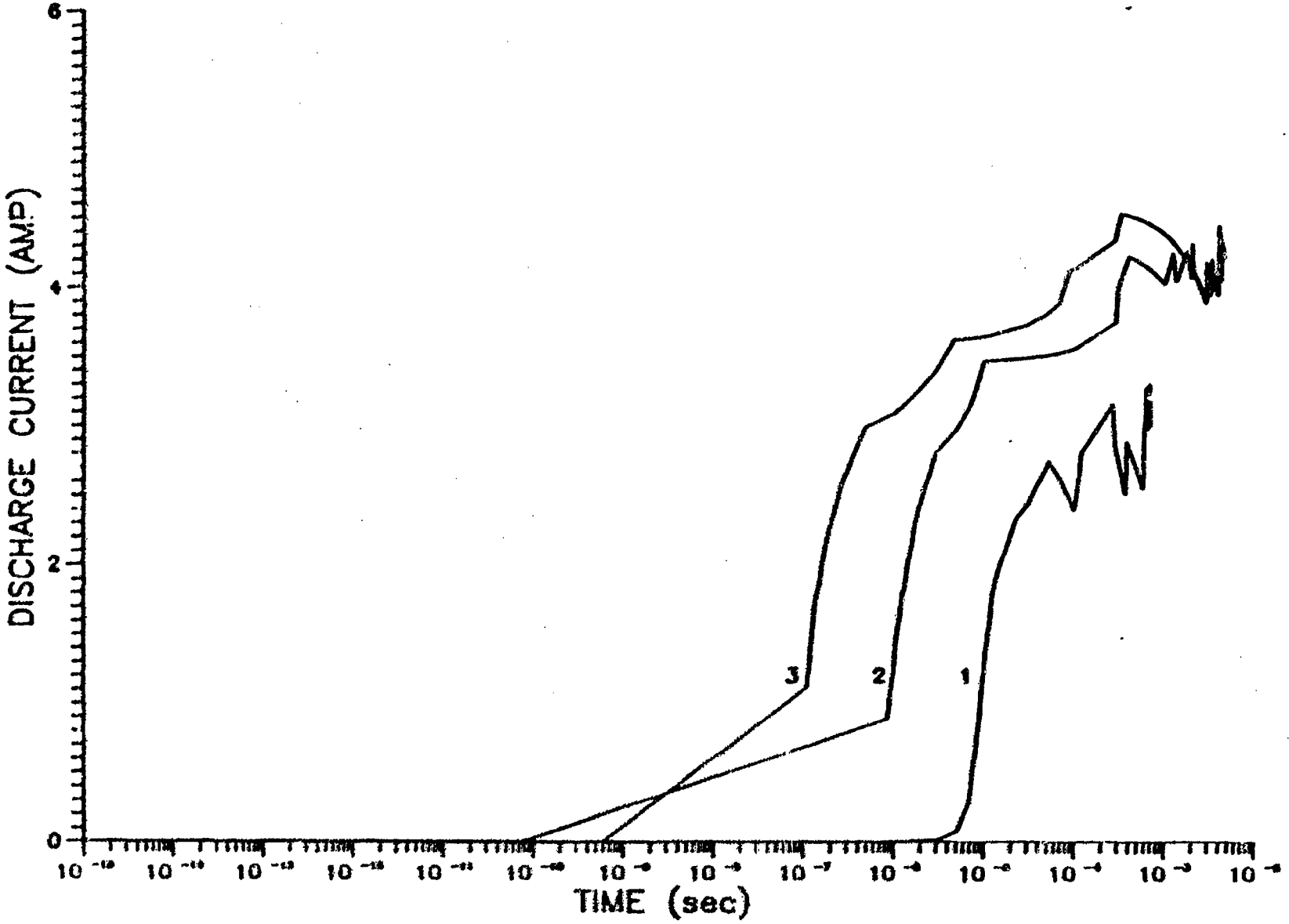
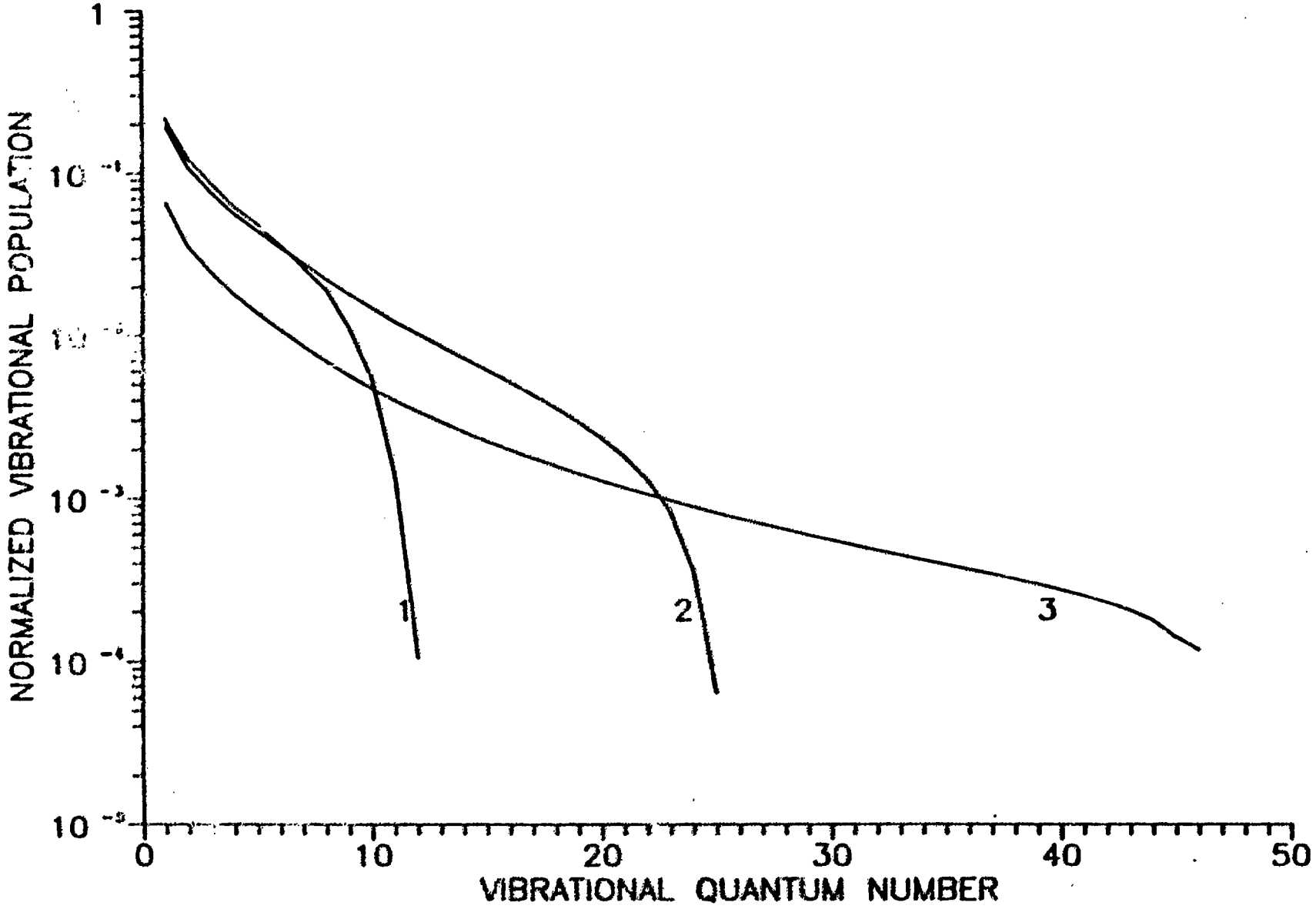
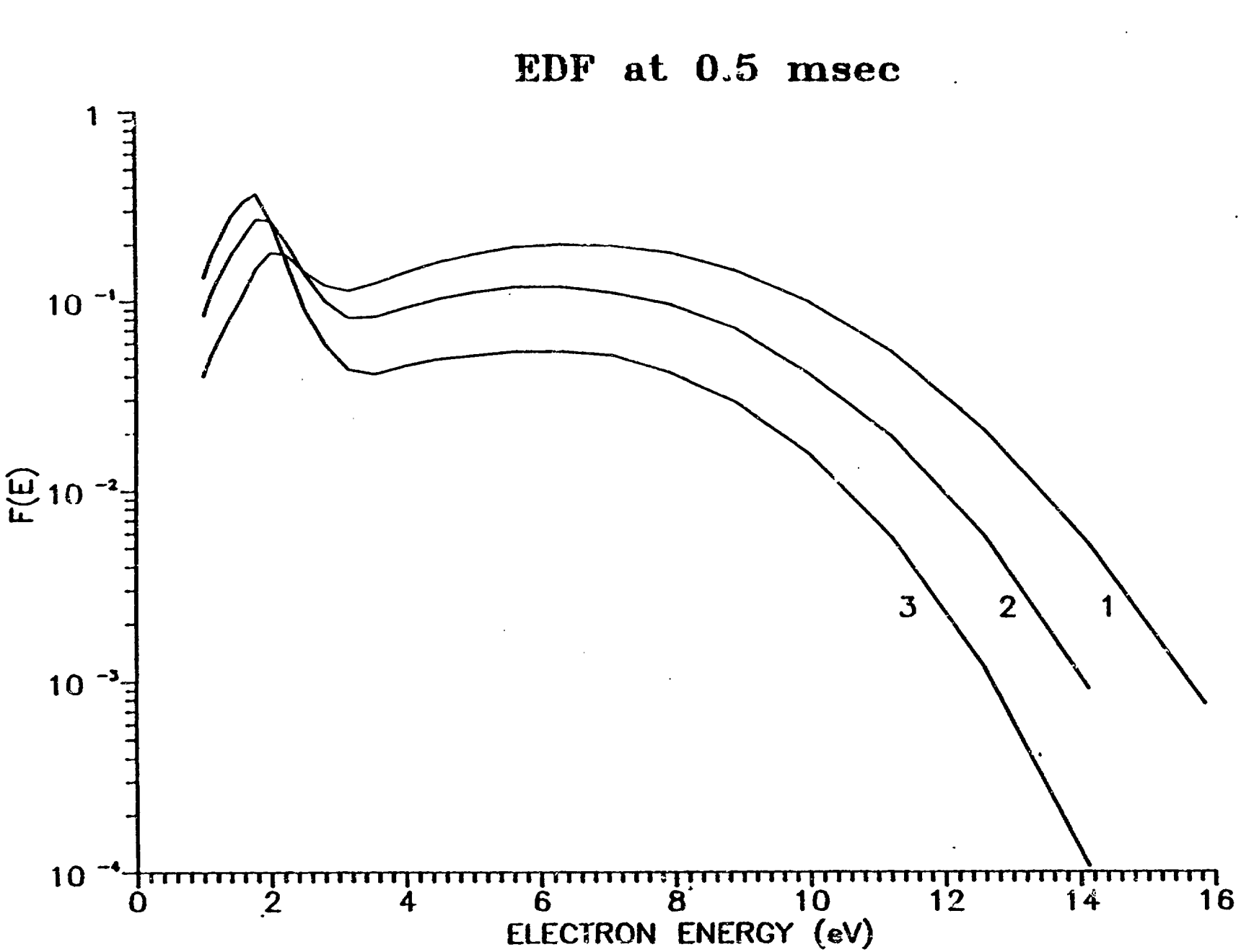


FIGURE 11

Vibrational population at 0.5 msec




EDF at 0.5 msec
1
10 -1
10 -2
10 -3
10 -4
F(E)
0
2
4
6
8
10
12
14
16
ELECTRON ENERGY (eV)
1
2
3

Time evolution of the electron density

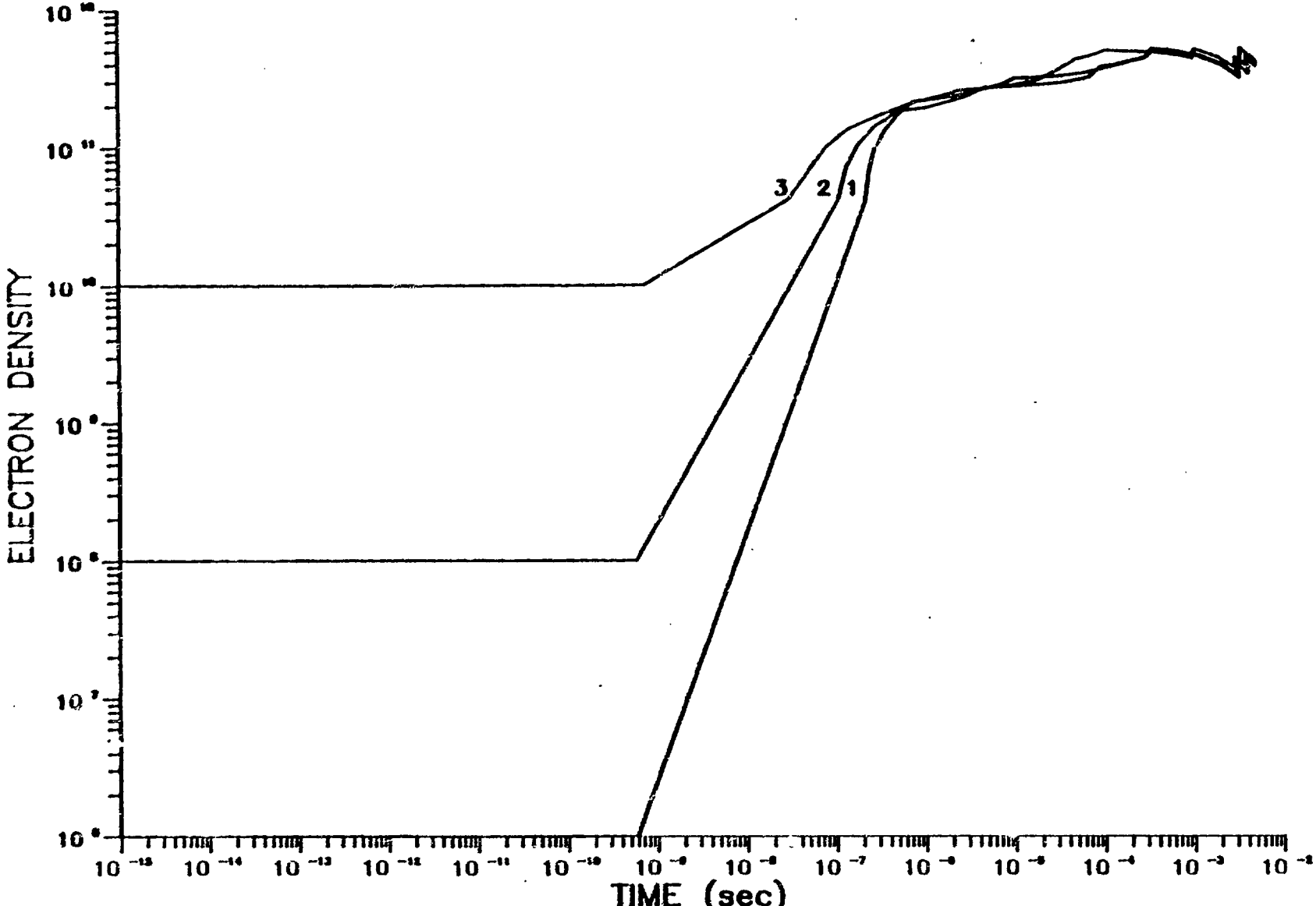
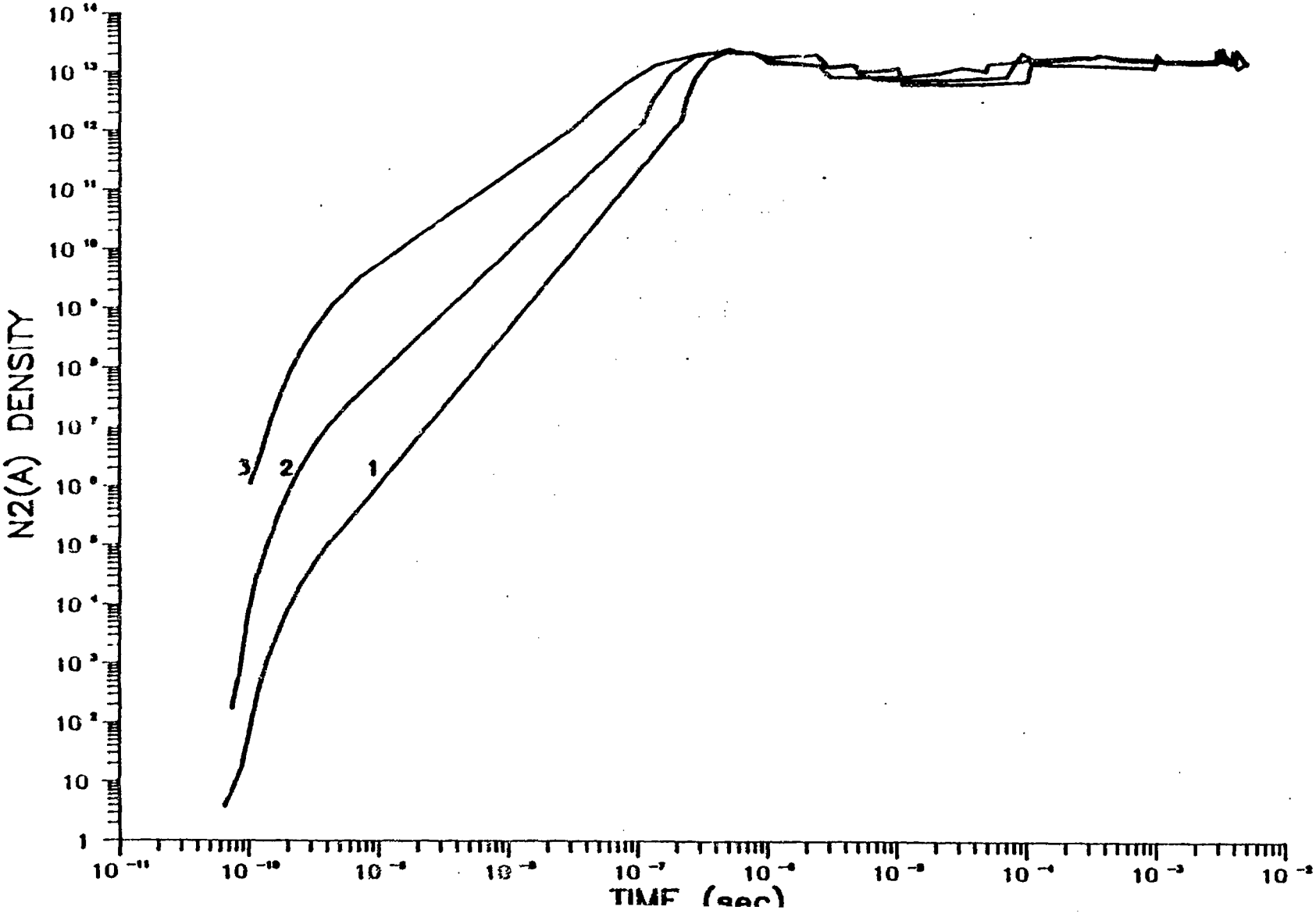


FIGURE 15

Time evolution of the nitrogen A-state density



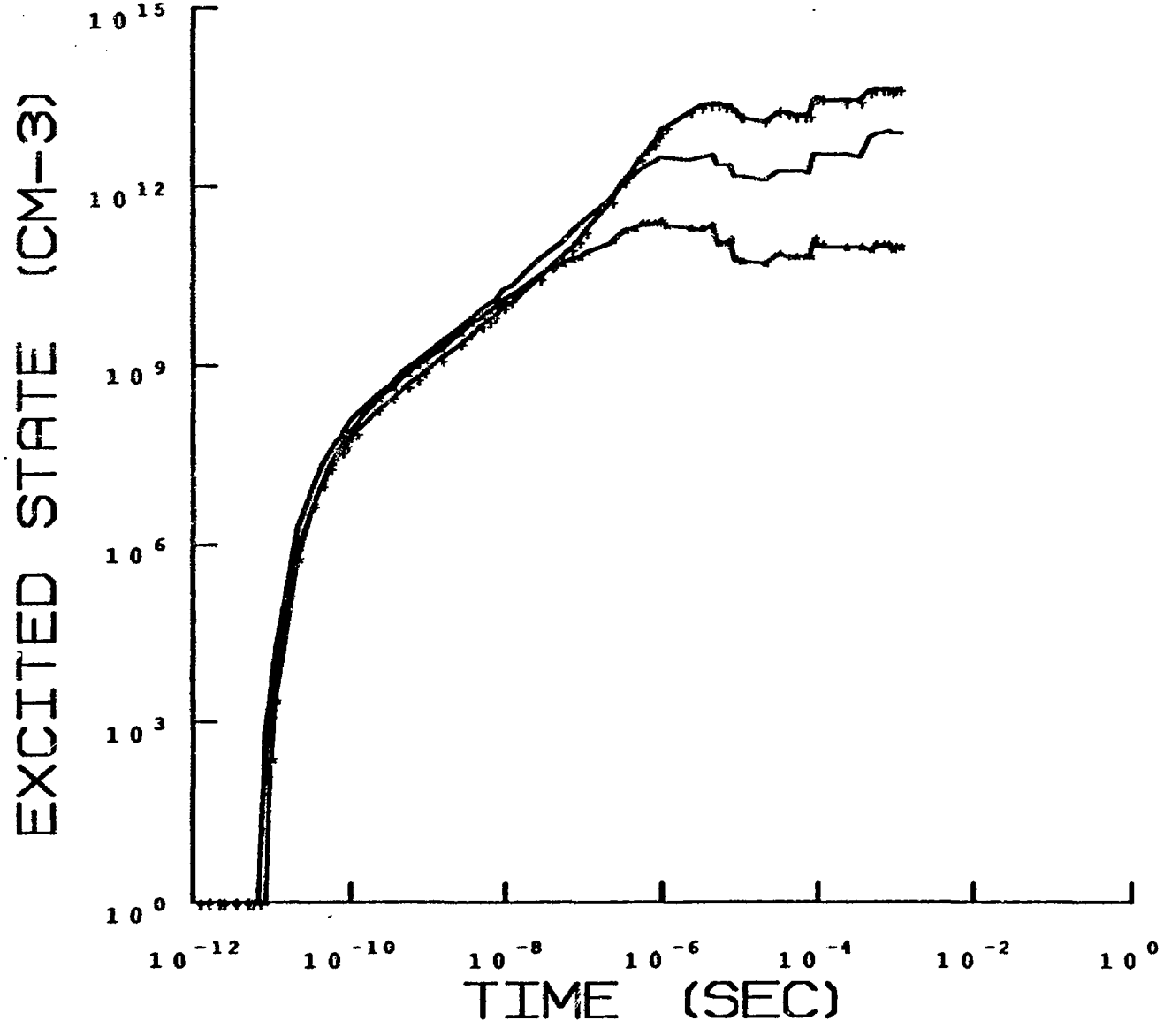
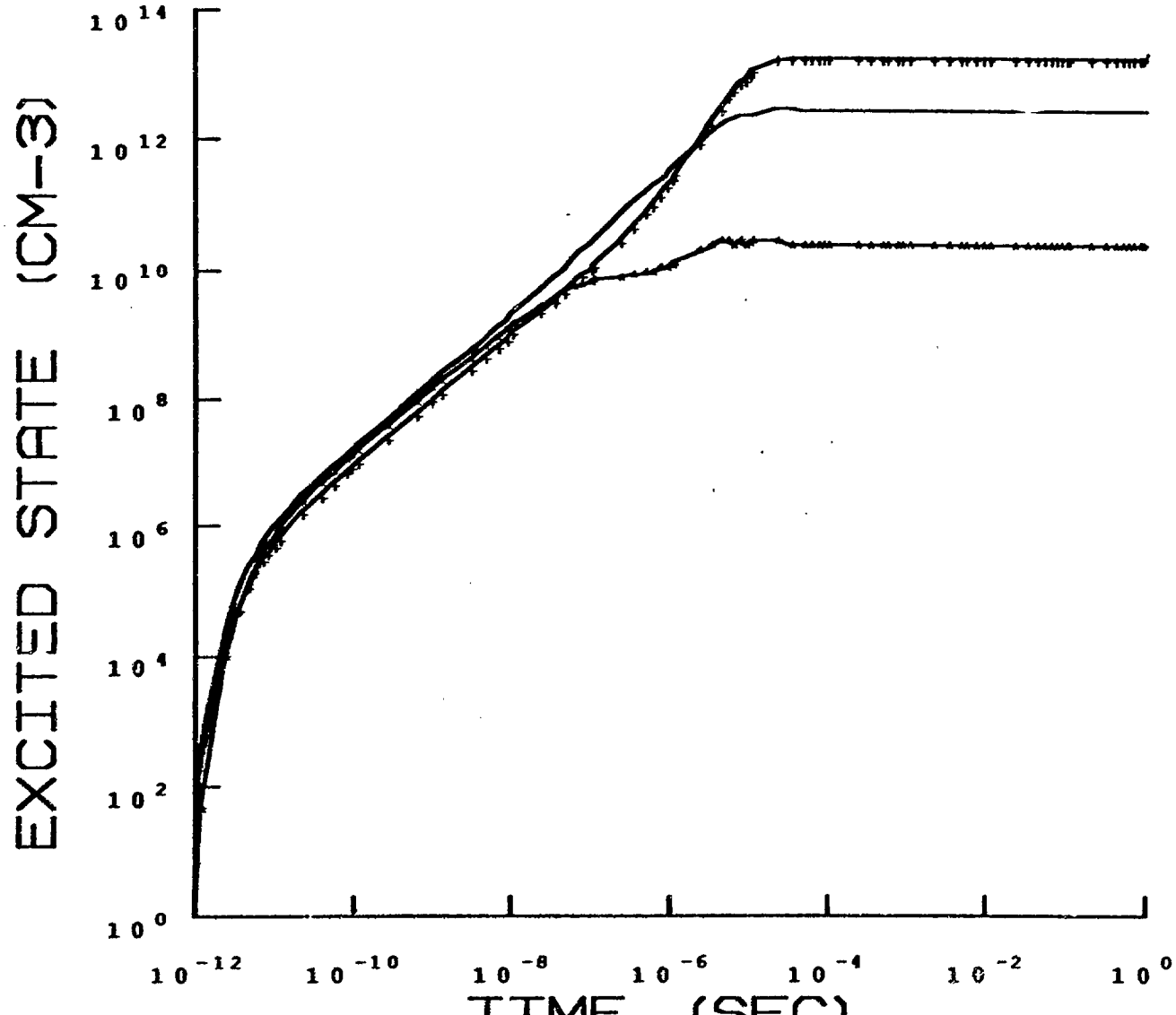

EXCITED STATE (CM-3)
10 15
10 12
10 9
10 6
10 3
10 0
10 -12
10 -10
10 -8
10 -6
10 -4
10 -2
10 0
TIME (SEC)

FIGURE 17


EXCITED STATE (CM-3)
10 14
10 12
10 10
10 8
10 6
10 4
10 2
10 0
10 -12
10 -10
10 -8
10 -6
10 -4
10 -2
10 0
TIME (SEC)

MODELING OF FAILURE MECHANISMS IN CERAMIC COMPOSITES UNDER FLEXURAL LOADING

by

L. R. Dharani

Department of Engineering Mechanics

University of Missouri-Rolla

Rolla, Missouri 65401

Contract No. F49620-85-C-0013/SB5851-0360

Purchase Order No. S-760-6MG-115

14 December, 1987

Final Report for period January 1987 - November 1987

SUMMARY

An analytical model is developed for predicting the fracture behavior of a fiber reinforced unidirectional ceramic matrix composite. In particular, the flexure test specimens used for strength and fracture toughness measurements is considered because of their common use and the prevailing desciepancy between the predicted and the observed failure stresses.

A combination of stress intensity factor and micromechanics approaches is employed in determining the critical crack length for various failure modes. The true stresses in the fibers bridging the matrix crack, compressive stresses in the beam and shear stresses at the crack tip are established as functions of the applied load, beam geometry, and the microstructural properties of the constituents. From the predicted stresses/stress intensity factor the failure modes are then established and summarized. The model predicts a transition in failure mode when the interface is strengthedned with fiber strength degradation.

INTRODUCTION

Ceramics, including glasses and glass-ceramics, form a technologically important class of materials. Their principal attributes are their low density, inertness and stability at high temparatures. However, because of their low tensile strengths and poor resistance to thermal and mechanical shock the use of ceramics in structural applications has been limited to situations where mainly compressive stresses are likely to be encountered. Ceramics can be used in high temperature structural applications provided a substantial improvement in their tensile and fracture toughness is achieved. Towards this end several methods of toughening for ceramics have been considered [1, 2] ; these are crack tip deflection and crack tip shielding which include fibrous toughening and transformation toughening. Of these methods the fibrous toughening has been shown to give the desired increase in fracture energy or toughness, as much as twenty times that of bulk ceramics. The high toughness of ceramic composites was first demonstrated using carbon fibers in glass and glass-ceramic matrices by Phillips and his co-workers [3 – 7]. More recently, the availability of continuous silicon carbide (SiC) fibers has led to the development of glass and glass-ceramic composites which are more resistant to high temperature oxidation than the carbon fibers composites.

Unlike the monolithic ceramic materials, the failure of ceramic matrix composites is very complex. Several investigators [3 – 13] have experimentally studied the failure of unidirectional ceramic matrix composites under tensile loading parallel to the aligned fibers. The brittle matrix is the first constituent to crack and this matrix crack which is normal to the applied stress passes through the central test section, and the applied load is supported entirely by the intact fibers bridging the crack. An increase in the applied stress, in most brittle matrix composites, results in the formation of multiple regularly spaced cracks in the central area. With further increase in the applied stress the crack deflects along the weak (frictional) fiber-matrix interface so that the fiber slips within the matrix. Debonding of the fiber from the matrix continues until the fiber fractures at a statistically weakest point whereupon the fiber is pulled out from the matrix. The brittle matrix undergoes extensive cracking

normal to the fibers, but the associated matrix cracking stress may be substantially greater than the catastrophic fracture stress of the monolithic ceramic.

Marshall and Evans [13], and Mah et al [14] observed that in room temperature four-point and three-point flexure tests the ultimate failure occurs either in compression or in shear but never in tension. The response was similar to that observed in the tensile test, except that the matrix cracks penetrated only to about midplane of the beam and the crack openings on the tensile surface did not become large enough to cause complete fiber pullout before compressive failure occurred in the opposite surface. High temperature tests on unidirectional ceramic matrix composites [15 – 16], however, do not reveal any multiple matrix cracking; only one (predominant) crack was observed in each specimen. Unlike in room temperature tests the high temperature tests showed no fiber bridging and a number of fibers at the crack site were broken. From the foregoing observations it is clear that the growth of an existing flaw in brittle matrix composites is not self-similar, and that the failure modes at room and high temperature could be significantly different.

The formation of matrix cracks, single or multiple, on the tensile side of the beam destroys the macroscopic uniformity of the beam and makes the use of flexure formula for bending stress calculations invalid. This fact is evident from the discrepancy between the values of the compressive stress calculated from the strain measurements [13] and those predicted by flexure formula. Inspite of this disadvantage, the flexure tests are still the most commonly used strength and fracture toughness tests because of the simplicity of the technique and the ease with which the specimens can be fabricated. Therefore, there is a need for understanding and characterizating the failure mechanisms in a brittle matrix composite under flexure, exhibiting single and multiple matrix cracks. In the present paper , an analytical model has been developed for the characterization of a unidirectional ceramic matrix composite under flexural loading with single and multiple matrix cracks on the tensile side. First, the stress intensity factor solutions for the single and multiple crack configurations are presented following the method of superposition [17]. Then, the true compressive stresses on the compression side of the beam, the stresses in the intact fibers bridging the matrix

crack and the shear stress at the matrix crack tip are determined. Finally , the modes of failure are investigated using an appropriate failure criterion.

FRACTURE MECHANICS ANALYSIS

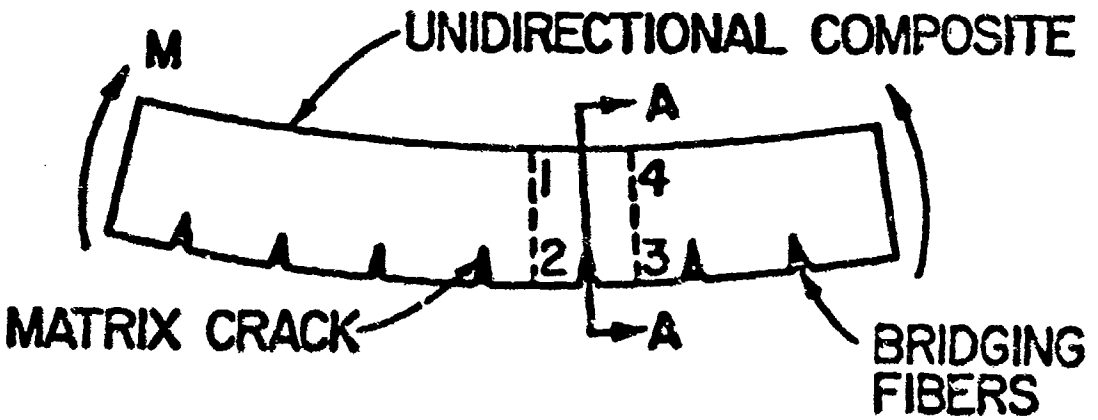
A schematic diagram of a four-point bend test specimen exhibiting multiple matrix cracks bridged by intact fibers is shown in Fig.1. For the sake of convenience the solution will be sought in two steps. The problem of a single matrix crack bridged by intact fibers in a long slender beam of unidirectional composite, with fibers parallel to the axis, will be considered first. Then, the solution will be modified to account for the multiple regularly spaced cracks on the tensile side of the beam. The beam is assumed to be under pure bending .

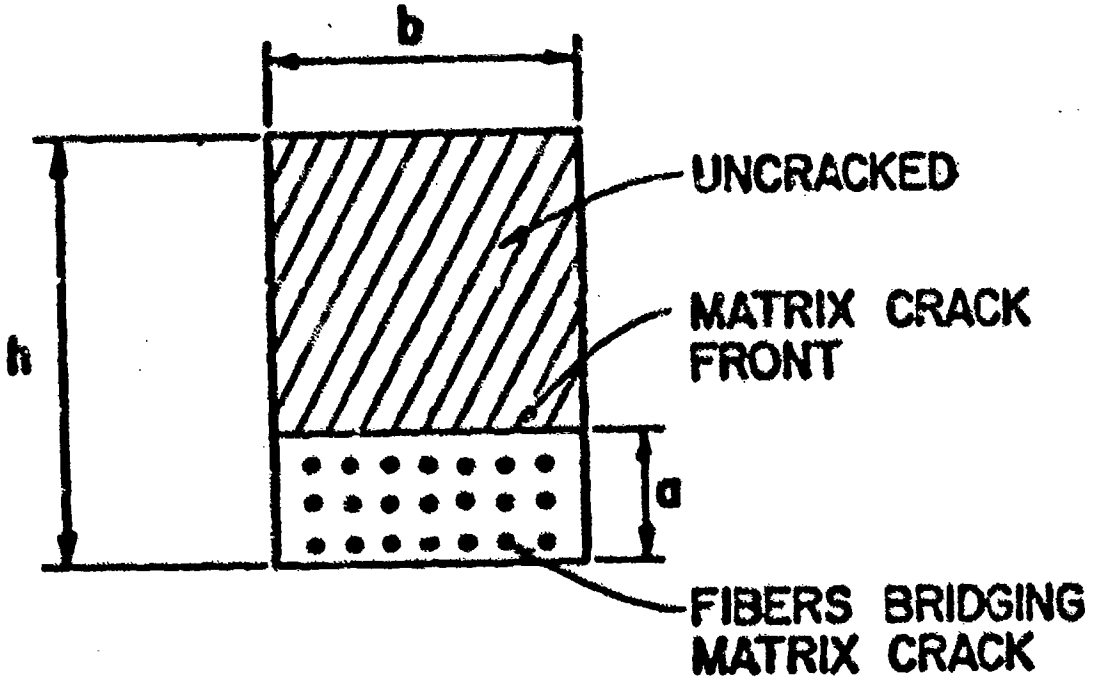
Single Matrix Crack

The stress intensity solution of a slender beam subjected to a bending moment, and containing a matrix crack bridged by fibers on the tension side will be obtained by the method of superposition, which was used by Marshall et al [17] to obtain the stress intensity factor solution for a unidirectional composite subjected to tensile load parallel to the axis of the fibers. With reference to Fig. 2, the stress intensity factor for the configuration (a) can be expressed as

$$K_I^{(a)} = K_I^{(b)} + K_I^{(c)} \tag{1}$$

where, $K_I^{(a)}$ is the desired stress intensity factor, and $K_I^{(b)}$ and $K_I^{(c)}$ are the stress intensity factors of configurations shown in Fig. 2(b) and 2(c), respectively. $K_I^{(b)}$ is the stress intensity factor of a homogenous beam containing a crack of length a on the tension side in which both the matrix and the fiber phases are broken. $w^{(b)}(x)$ in Fig. 2(b) is the corresponding crack opening displacement . It will be assumed that there is no relative displacement between the fibers and the matrix in Fig. 2(b). $K_I^{(c)}$ is the stress intensity factor of a long finite width strip containing an edge notch subjected to a symmetric traction $p^{(c)}(x)$ on the crack faces. The traction $p^{(c)}(x)$ is chosen such that the fiber ends may displace relative to the matrix and be rejoined along the x-axis. During the process of rejoining the fiber-matrix interface may fail and slip over some length and





SECTION A-A

Figure 1. A Unidirectional Composite Beam Exhibiting Multiple Matrix Cracking.

remain bonded over the remainder . This is accounted for while determining the traction $p^c(x)$ from the mechanics of fiber pullout.

In general, by means of weight function method [18] , the stress intensity factor and the crack opening displacement of a finite width strip with single edge crack, subjected to a distributed force q(x) on the crack surface may be given as

$$K_I = 2\int_0^a q(x)m(x,a)dx \tag{2}$$

$$u(x) = \int_x^a \frac{2}{E'} K_I m(x,\eta)d\eta \tag{3}$$

where $E' = E$ for plane stress and $E' = E/(1-\nu^2)$ for plane strain, u(x) is the crack opening displacement and m(x,a) is the weight function given by

$$m(x,a) = \frac{1 + m_1(a)(\frac{a-x}{a}) + m_2(a)(\frac{a-x}{a})^2}{\sqrt{2\pi(a-x)}} \tag{4}$$

with

$$m_1(a) = 0.6147 + 17.1844(\frac{a}{h})^2 + 8.7822(\frac{a}{h})^6$$

$$m_2(a) = 0.2502 + 3.2899(\frac{a}{h})^2 + 70.0444(\frac{a}{h})^6$$

By setting $q(x) = \sigma_0(1 - 2x/h)$, the stress intensity factor , and the displacement for the configuration shown in Fig. 2(b), subject to a remote bending moment, M , can be obtained from eqns (2) and (3) as

$$K_I^{(b)} = \sigma_0\sqrt{\frac{2a}{\pi}}\int_0^1 (1 - \frac{2a}{h}t)\frac{w(a,t)}{\sqrt{1-t}}dt \tag{5}$$

$$u^{(b)}(t) = \frac{2a}{\pi E'}\sigma_0\int_t^1 \frac{w(a\xi,\frac{t}{\xi})}{\sqrt{\xi - t}}\int_0^t (1 - \frac{2a}{h}\eta)\frac{w(a\xi,\frac{\eta}{\xi})}{\sqrt{\xi - \eta}}d\eta d\xi \tag{6}$$

where $\sigma_0 = \frac{6M}{bh^2}$, $t = x/a$ and $w(a,t) = 1 + m_1(a)(1-t) + m_2(a)(1-t)^2$. In order to evaluate the stress intensity factor $K_I^{(b)}$ for the configuration Fig. 2(c) , it is necessary to

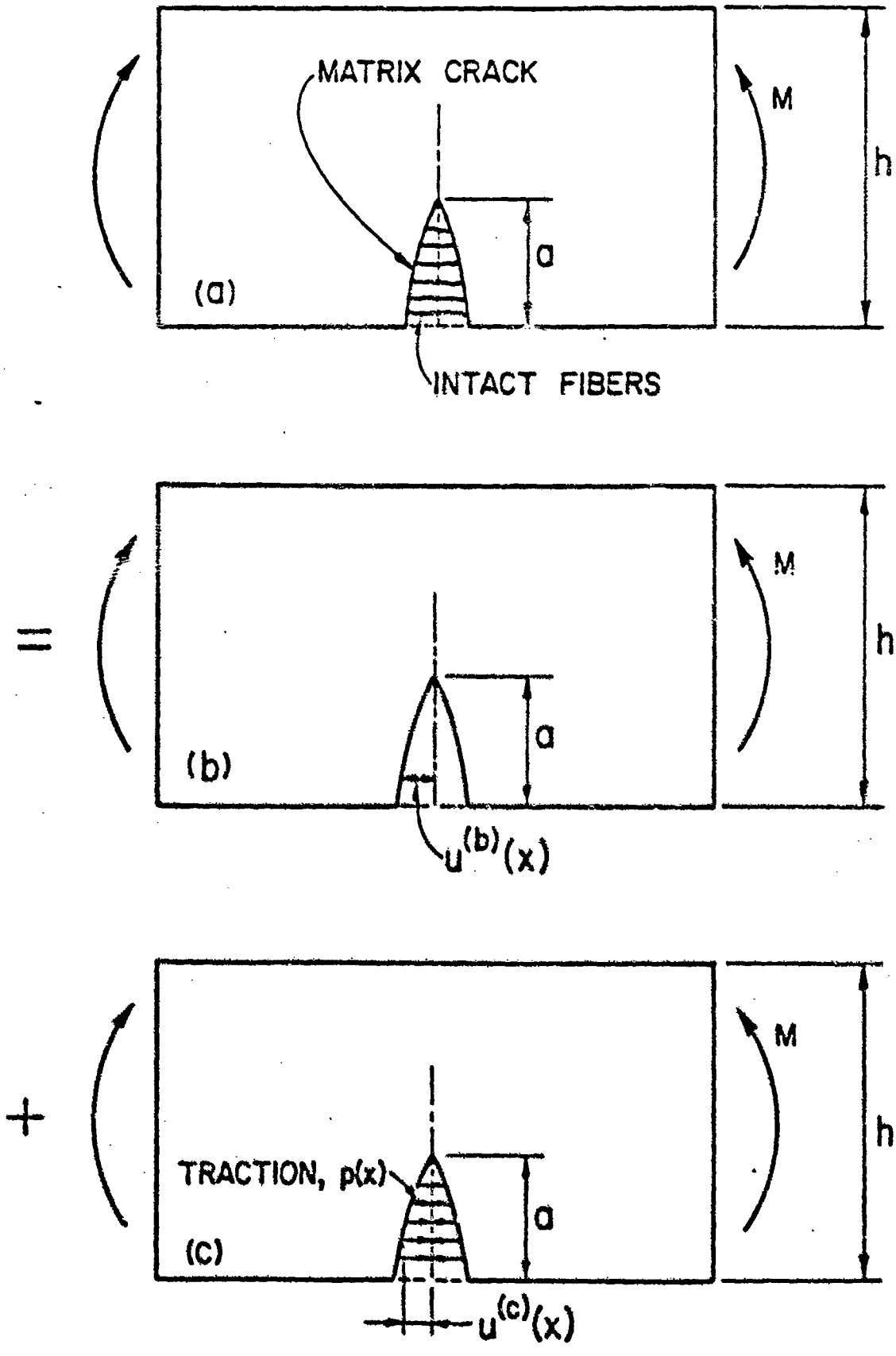


Figure 2. Superposition of Stress Factor Solutions.

calculate the pressure distribution $p^{(c)}(x)$ seperately. For a unidirectional composite with a weakly bonded interface, such as frictional, the closure pressure can be related to microstructural properties of the constituents, and the fiber displacement by considering the process of fiber pullout [17] as

$$p^{(c)}(x) = 2V_f[E_f\tau(1+\eta)\frac{u^{(c)}(x)}{R}]^{1/2} \tag{7}$$

where E_f and V_f are the elastic modulus and the volume fraction of the fiber , τ is the frictional stress at the fiber-matrix interface, $\eta = E_fV_f/E_mV_m$, and R is the fiber radus. The stress intensity factor for the configuration in Fig. 2(c) is then given by

$$K_I^{(c)} = \sigma_0\sqrt{\frac{2a}{\pi}}\int_0^1 p^{(c)}(t)\frac{w(a,t)}{\sqrt{1-t}}dt \tag{8}$$

$$u^{(c)}(t) = \frac{2a}{\pi E'}\sigma_0\int_t^1 \frac{w(a\xi,\frac{t}{\xi})}{\sqrt{\xi-t}}\int_0^\xi p^{(c)}(\eta)\frac{w(a\xi,\frac{\eta}{\xi})}{\sqrt{\xi-\eta}}d\eta d\xi \tag{9}$$

After substituting into eqn (1) with considerable manipulation the final solution can be given as

$$K_I^{(a)} = \sigma_0\sqrt{\frac{2a}{\pi}}\int_0^1 (1-\frac{2a}{h}t-\sqrt{\bar{u}(t)})\frac{w(a,t)}{\sqrt{1-t}}dt \tag{10}$$

$$\bar{u}(t) = \frac{2a}{\pi E'u_0}\sigma_0\int_t^1 \frac{w(a\xi,\frac{t}{\xi})}{\sqrt{\xi-t}}\int_0^\xi (1-\frac{2a}{h}\eta-\sqrt{\bar{u}(\eta)})\frac{w(a\xi,\frac{\eta}{\xi})}{\sqrt{\xi-\eta}}d\eta d\xi \tag{11}$$

where $\bar{u}(x)$ is the normalized crack opening displacement given by

$$\bar{u}(x) = \frac{u^{(a)}(x)}{u_0} = 4u^{(a)}(x)V_f^2(1+\eta)\tau E_f/(\sigma_0^2R) \tag{12}$$

Multiple Matrix Cracks

In this section an approximate stress intensity factor solution for the beam configuration, with multiple regularly spaced matrix cracks bridged by intact fibers, is presented. Consider a portion of the beam containing one of the matrix cracks, enclosed by sections 1-2 and 3-4 as shown in Fig.1. Both sections are at a distance equal to one-half the crack spacing from the enclosed crack, so that the distance between 1-2 and 3-4 is exactly equal to the crack spacing. A comparasion of the element 1-2-3-4 of Fig.1 with Fig.2 suggests that the solutions to the multiple crack problem and the single crack problem must be of the same form. The difference in the solution is due to the fact that the distributed end load with zero resultant force and resultant moment equal to M is applied at a finite distance from the plane of the crack in the case of multiple cracking whereas the same is applied at infinity in the case of single crack. With this observation the stress intensity factor $K_I^{(a,m)}$ of the beam with multiple regularly spaced cracks can be written as

$$K_I^{(a,m)} = K_0 F_1(a,h) F_2(a,c) \tag{13}$$

where K_0 is the stress intensity factor of a half plane with single edge crack, $F_1(a,h)$ and $F_2(a,c)$ are the correction factor due to finite width h and multiple cracks with spacing of 2c. Equation (13) implies that the effects of finite width of the beam and multiple cracks on the stress intensity factor are independent.

For a half-plane with a row of parallel edge cracks of depth a and spacing 2c, subjected to a uniform stress σ_0 on the crack surface, the stress in the vicinity of crack tip is given by [19]

$$\sigma_y = k(a,c)\sigma_0\left(\frac{ac}{a+c}\right)^{1/2}\frac{1}{\sqrt{2r}}, \qquad r \to 0 \tag{14}$$

where

$$k(a,c) = \frac{1}{\sqrt{\pi}}\left[1 + \frac{c}{2d} + \frac{3c^2}{8d^2} + \frac{5c^3}{16d^3} + \frac{35c^4}{128d^4} + \frac{63c^5}{256d^5} + \frac{231c^6}{1024d^6}\right] + 22.501\frac{c^7}{d^7} - 63.502\frac{c^8}{d^8} + 58.045\frac{c^9}{d^9} - 17.577\frac{c^{10}}{d^{10}} \tag{15}$$

with d = a + c, so that, the stress intensity factor K_I is given by

$$K_I = \lim_{r\to 0}\sqrt{2\pi r}\,\sigma_y = \sigma_0\sqrt{\pi}\,\sqrt{\frac{ac}{a+c}}\,k(a,c) \tag{16}$$

By comparing eqn (16) with the stress intensity factor for a single edge crack in a half-plane subjected to a remote normal stress σ_0, we obtain

$$F_2(a,c) = 0.891\sqrt{\frac{c}{a+c}}\,k(a,c) \tag{17}$$

Finally, following the procedure of single crack problem, the stress intensity factor and the crack opening displacement for the problem of multiple matrix cracks are obtained as

$$K_I^{(a,m)} = 0.7111\sigma_0\sqrt{\frac{ac}{(a+c)}}\,k(a,c)\int_0^1\left[1-\frac{2a}{h}t-\sqrt{\bar{u}(t)}\right]\frac{w(a,t)}{\sqrt{1-t}}dt \tag{18}$$

$$\bar{u}(t) = \frac{2a\sigma_0}{1.122\pi E'u_0}\int_t^1 k(a\xi,c)\sqrt{\frac{c}{a}\xi+c}\,\frac{w(a\xi,\frac{t}{\xi})}{\sqrt{\xi-t}}\times$$

$$\int_0^\xi\left(1-\frac{2a}{h}\eta-\sqrt{\bar{u}(\eta)}\right)\frac{w(a\xi,\frac{\eta}{\xi})}{\sqrt{\xi-\eta}}d\eta d\xi \tag{19}$$

NUMERICAL PROCEDURE

The governing equations of both single crack and multiple crack problems represent a set of coupled singular integral equations for which a closed form solution cannot be readily obtained. Therefore, a numerical technique is developed to solve these equations. This technique is illustrated in this section by applying it to the solution of multiple crack problem.

The singularities in the integral eqns (18) and (19) can be removed by setting $t = \sin^2\theta$, $\xi = t\cosh^2\phi$, and $\eta = \sin^2\theta$. This change in variables reduces eqns (18) and (19) to

$$K_I^{(a,m)} = 1.4223\sigma_0\sqrt{\frac{ac}{a+c}}\,k(a,c)\int_0^{\frac{\pi}{2}}\left[1-\frac{2a}{h}\sin^2\theta-\sqrt{\bar{u}(a\sin^2\theta)}\,\right]w(a,\sin^2\theta)\sin\theta\,d\theta \quad (20)$$

$$\bar{u}(t) = \gamma t\int_0^{\cosh^{-1}1/\sqrt{t}}\sqrt{\frac{c}{a\cosh^2\phi+c}}\,k(a\cosh^2\phi,c)w(a\cosh^2\phi,\frac{1}{\cosh^2\phi})\cosh^2\phi\times$$

$$\int_0^{\frac{\pi}{2}}\left[1-\frac{2a}{h}t\cosh^2\phi\sin^2\theta-\sqrt{\bar{u}(a\cosh^2\phi\sin^2\theta)}\,\right]w(a\cosh^2\phi,\sin^2\theta)\sin\theta\,d\theta\,d\phi \quad (21)$$

where

$$\gamma = \frac{8\sigma_0 a}{1.122\pi E u_3}$$

The displacement $\bar{u}$ is represented in the form

$$\bar{u}(x) = \sum_{i=1}^{n} u_i\psi_i(x) \quad (22)$$

where $\bar{u}_i$ are the values of $\bar{u}$ at Gauss-points, and $\psi_i(x)$ are the shape functions, such as linear, quadratic or cubic spline. Substituting eqn (22) into (21) and then satisfying the resulting equation at the n Gauss points yields the following system of non-linear equations for the unknowns $\bar{u}_i$

$$f(t_i) = \gamma t_i\int_0^{\cosh^{-1}1/\sqrt{t_i}}\sqrt{\frac{c}{at_i\cosh^2\phi+c}}\,k(at_i\cosh^2\phi,c)w(at_i\cosh^2\phi,\frac{1}{\cosh^2\phi})\times$$
$$\cosh^2\phi\int_0^{\frac{\pi}{2}}\left[1-\frac{2a}{h}t_i\cosh^2\phi\sin^2\theta-\sqrt{\bar{u}(at_i\cosh^2\phi\sin^2\theta)}\,\right]w(at_i\cosh^2\phi,\sin^2\theta)\times$$

$$\sin\theta\,d\theta\,d\phi-\bar{u}(t_i)=0 \qquad i=1,2,\ldots,n \quad (23)$$

These equations are solved numerically, and then the value of $K_I^{(a,m)}$ is obtained by evaluating the integral in eqn (20) by Gauss quadrature.

To check the accuracy and the speed of convergence of the solution linear, cubic polynorminal, cubic spline and Lagrange polynorminal shape functions were tried with different numbers of Gauss quadrature points. The linear and the cubic spline shape functions gave the optimum results in terms of the computation time and accuracy.

The technique seems to work for the range of γ values covering most ceramic matrix composite systems .

FAILURE MODES

For a unidirectional composite under flexure with single or multiple matrix cracking on the tension side the following failure modes are possible

(1) Self-similar extension of a bridged matrix crack

(2) Delamination parallel to the axis of the beam

(3) Fracture of the bridging fibers

(4) Compression failure (crushing) of the composite

Each of these failure modes would be governed by a different failure criterion, as described below.

Self-Similar Matrix Cracking

This mode of failure is assumed to occur when the stress intensity factor K_m of the matrix exceeds its critical stress intensity factor of the monolithic matrix material, K_m^c. In the foregoing sections the stress intensity factor for the unidirectional composite is evaluated. Following the procedure of Marshall et al [17], in which it is assumed that the matrix and composite stress intensities scale with the stresses, the stress intensity factor for the matrix can be written as

$$K_m = K_I E'/E_m \tag{24}$$

where K_I is the composite stress intensity factor, given by either eqn (10) or eqn (18). The use of such a rule of mixture type of relation may be justified, based on a recent investigation by Chiang [20], as long as the crack length is at least three times fiber diameter.

Delamination

The delamination failure mode must be distinguished from the fiber-matrix interface debonding in that this occurs in the form of splitting of the matrix material parallel to the fibers. The matrix splitting is assumed to occur when the maximum shear stress at the crack tip exceeds the matrix ultimate shear stress. By using the shear lag theory [21], the maximum shear stress in the matrix is estimated to be

$$\tau_{max} = 2G_m u_n / H \tag{25}$$

where G_m is the matrix shear modulus, H is the fiber center line spacing and $2u_n$ is the crack opening displacement at the first fiber behind the crack tip. For delamination parallel to the fiber axis the maximum shear stress must exceed the ultimate shear strength of the matrix, τ_m^u.

Fiber Failure

In the plane of the matrix crack , the bridging fibers carry an additional load and as a result the fibers may fail. Even though the location of the actual failure must be in the plane of the matrix crack, because of the statistical nature of the fiber strength this could be random. The maximum stress in the intact fiber in the plane of the matrix crack is

$$\sigma_f^{max} = \sigma_0(1 + \sqrt{\bar{u}(x)})|_{x=0} \tag{26}$$

The fiber breaks occur when the maximum fiber stress exceeds the fiber ultimate tensile strength σ_f^u.

Compressive Failure

Because of the cracking in the tensile side of the beam the apparent shift of neutral axis will result in compressive stresses being different from those predicted by flexure formula. The compressive failure of the material may occur in the form of crushing. An estimate of the true compressive stress for the beam with a single or multiple matrix cracks is carried out.

To evaluate the true stress on the compression side of the beam, it is assumed that the stress in the uncracked region varies linearly except in the vicinity of crack tip, at which it has a singularity, as shown in Fig.3 :

$$\sigma(x) = \begin{cases} \sigma_0(1 + \sqrt{\bar{u}(x)}) & 0 \le x \le a \\ K_I/\sqrt{2\pi x} & a \le x \le a + x_1 \\ \dfrac{\sigma_1 + \sigma_2}{h - a - x_0}(h - x) - \sigma_2 & a + x_1 \le x \le h \end{cases} \tag{27}$$

where σ_1, σ_2, x_1 are determined from the following equilibrium and continuity conditions

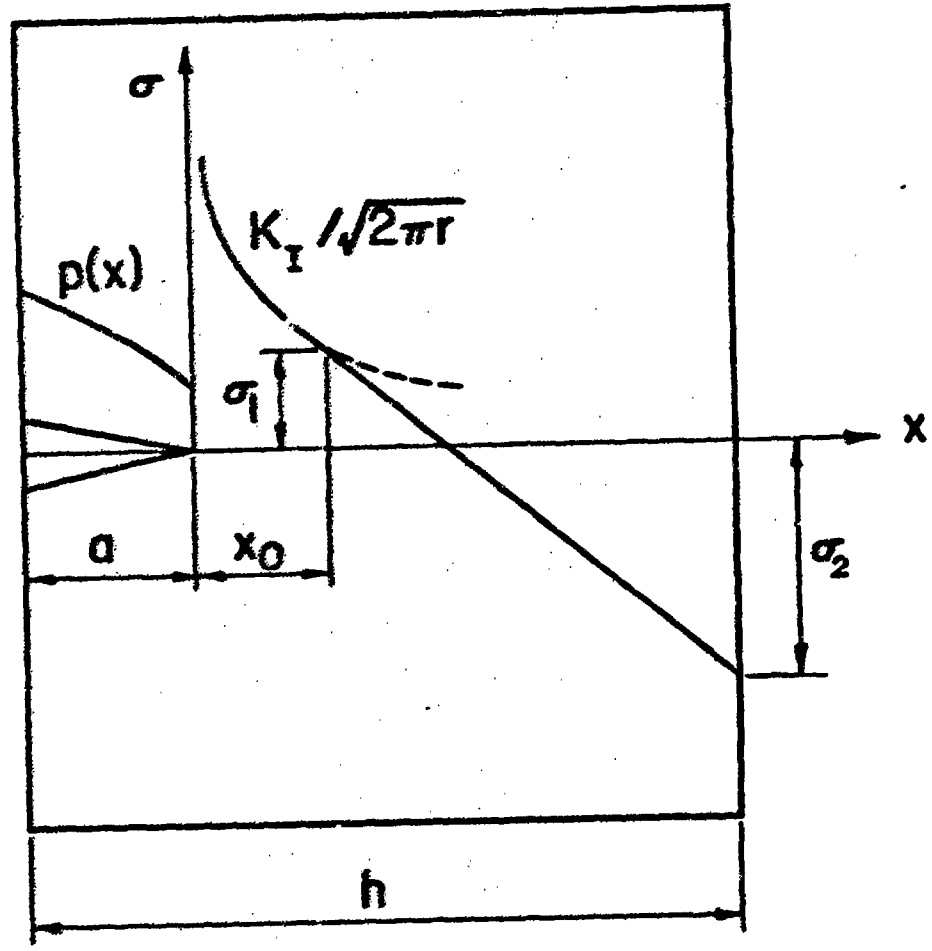


Figure 3. A Representative Stress Distribution at a Matrix Crack Location.

$$\sigma_0\left[\frac{a}{h}(h-a)-u_I\right]+1/2(\sigma_1-\sigma_2)(h-a-x_1)=0 \tag{28}$$

$$\sigma_0\left(a^2\left(\frac{1}{2}-\frac{2a}{3h}\right)-u_{II}\right)+2\sigma_1 x_1\left(a+\frac{x_1}{3}\right)+$$

$$\frac{1}{6}\left(h^2-(a+x_1)^2\left(2\sigma_1-\sigma_2-(\sigma_1+\sigma_2)\frac{h}{h-a-x_1}\right)\right)=M \tag{29}$$

$$\frac{K_I}{\sqrt{2\pi x_1}}=\sigma_1 \tag{30}$$

where

$$u_I=\int_0^a\sqrt{\bar{u}(x)}\,dx$$

$$u_{II}=\int_0^a x\sqrt{\bar{u}(x)}\,dx$$

Eqns (28-30) are solved numerically. For compressive failure the maximum compressive stress σ_2 must exceed the ultimate compressive stress of the composite.

RESULTS AND DISCUSSION

In order to illustrate the capabilities of the analytical model developed here, we present some representative results for a typical ceramic matrix composite system such as SiC/LAS with E_f = 200 GPa , E_m = 85 GPa , ν_f = 0.5 and K_c = 2 $MPa\ m^{1/2}$. To study the effect of interface strength on failure modes a range of values for interface strength τ from 0.5 to 30 MPa are considered. For a similar reason various fiber sizes are also considered, R = 8 - 150 microns. Since it has been observed that fiber strength degrades considerably when the composite is exposed to elevated temperatures we will present the results for two fiber strengths, that is , σ_f = 1450 MPa and 870 MPa . The lower value represents a 40% degradation in fiber strength at elevated temperature . The higher value of τ also represents the fiber-matrix interface strength that has been observed in experiments by Luh and Evans [15].

The effect of varying the interface strength and the fiber size on the critical crack length is presented in Fig. 4 and 5. It can be seen from Fig. 4 that a stronger interface results in a higher critical stress for a self-similar crack extension at a given crack

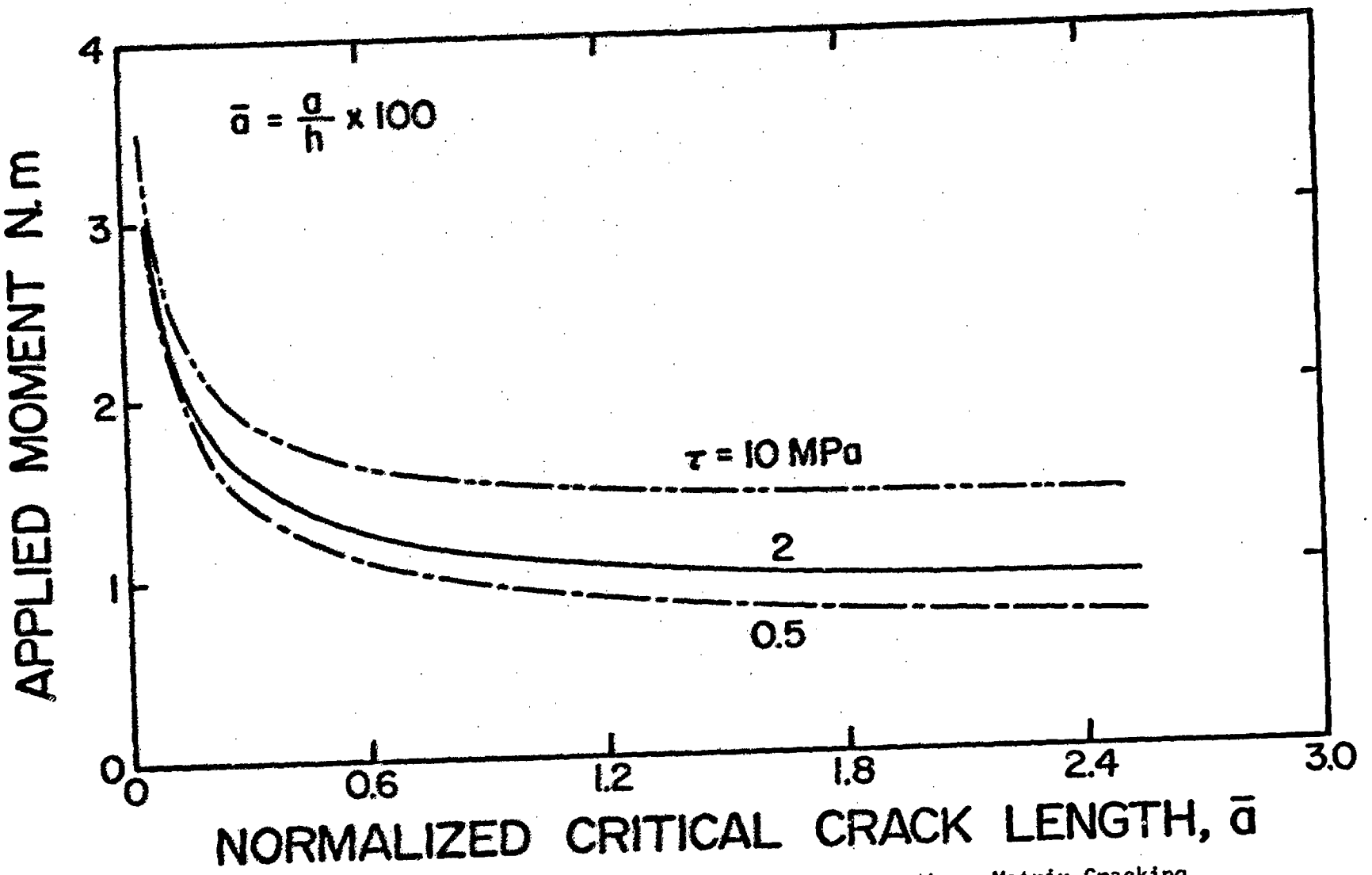


Figure 4. Effect of Interface Strength on Matrix Cracking.

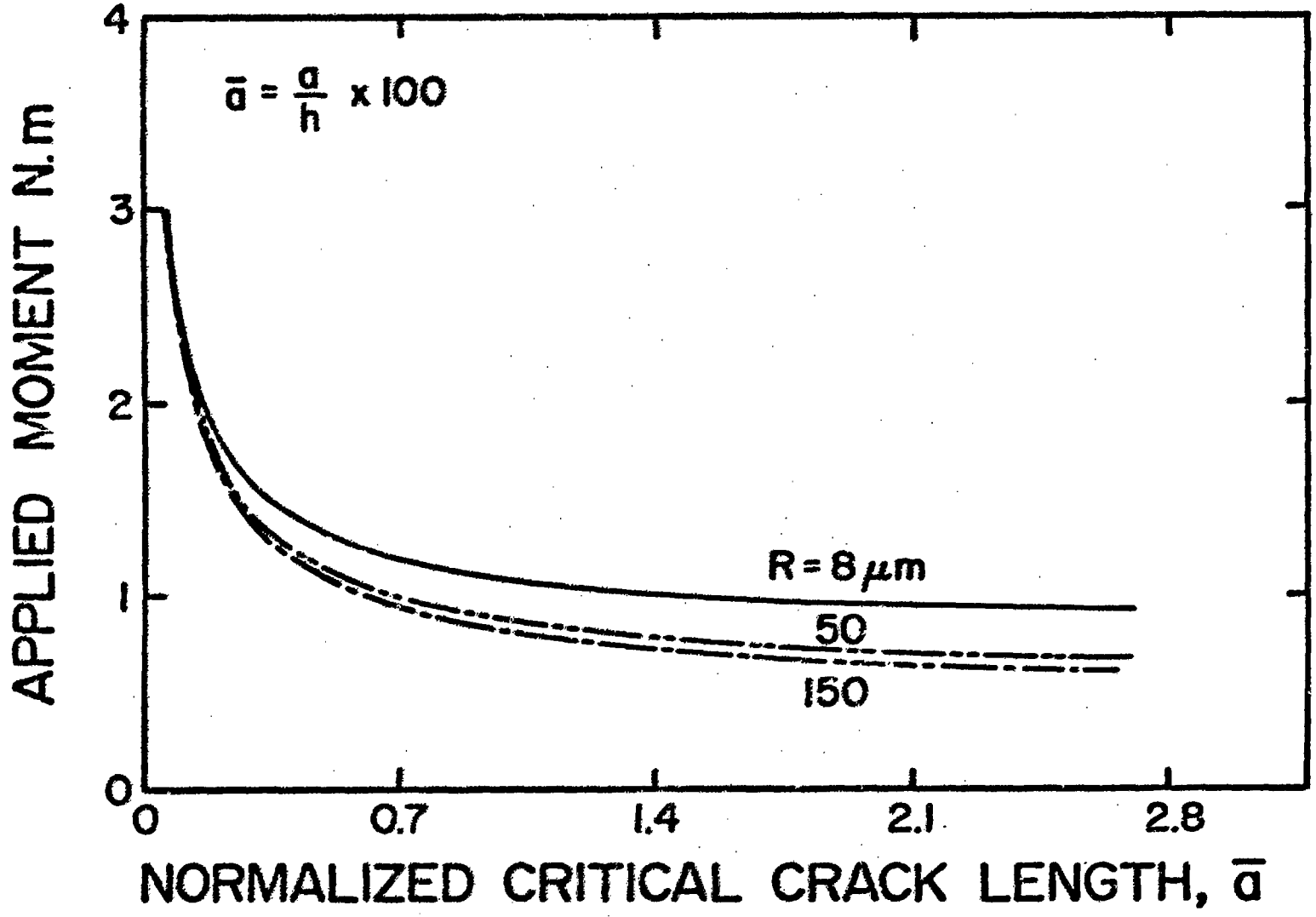


Figure 5. Effect of Fiber Size on Matrix Cracking.

length. This, however, does not imply a higher fracture toughness for the composite, because as the crack grows in a self-similar fashion the mode of failure may change to that of fiber fracture and result in a catastrophic failure of the composite. On the other hand, when the diameter of the fiber increases, the critical stress will decrease for a given interface strength which could also be seen from considering the equilibrium of the fiber.

Fig. 6 and 7 show the relationship between the normalized maximum stress in the first bridging fiber, $\bar{\sigma}_{f_{max}}$, the maximum shear stress in the matrix at the crack tip, $\bar{\tau}_{max}$, and the stress intensity factor of the matrix, $\bar{K}_i$, and the applied moment, for two crack sizes; small crack, a = 0.025h , and large crack, a = 0.25h. The normalization is being carried out with respect to the corresponding critical values as follows :

$$\bar{\tau} = \tau_{max}/\tau_m^*$$

$$\bar{\sigma}_f = \sigma_{f_{max}}/\sigma_f^*$$

$$\bar{K}_i = K_i/K_m^*$$

The failure of the composite would then occur when any of these normazed variables attain a value of 1.0. For each crack size the following two cases are considered :

(a) τ = 2 MPa and σ_f^* = 1450 MPa which represents the typical room temperature properties for a ceramic fiber/glass-ceramic matrix composite, such as SiC/LAS.

(b) τ = 30 MPa and σ_f^* = 870 MPa which represents the composite in case (a) at elevated temperature in which interface strengthening and fiber strength degradation have taken place.

For relatively small cracks such as those in the range of inherent flaw size, Fig. 6, the current model predicts the failure to occur in a self similar-fashion, that is, the matrix crack bridged by intact fibers. In this particular choice of material properties the self-similar extension and matrix splitting (debonding) seem to occur simultaneouly which is a mere coincidence. If the fiber-matrix interface is strengthened along with a 40% decrease in fiber strength, the failure mode would be that of fiber fracture. For large cracks, the room temperature failure would be by debonding and the difference in the critical stresses between the self-similar crack extension and

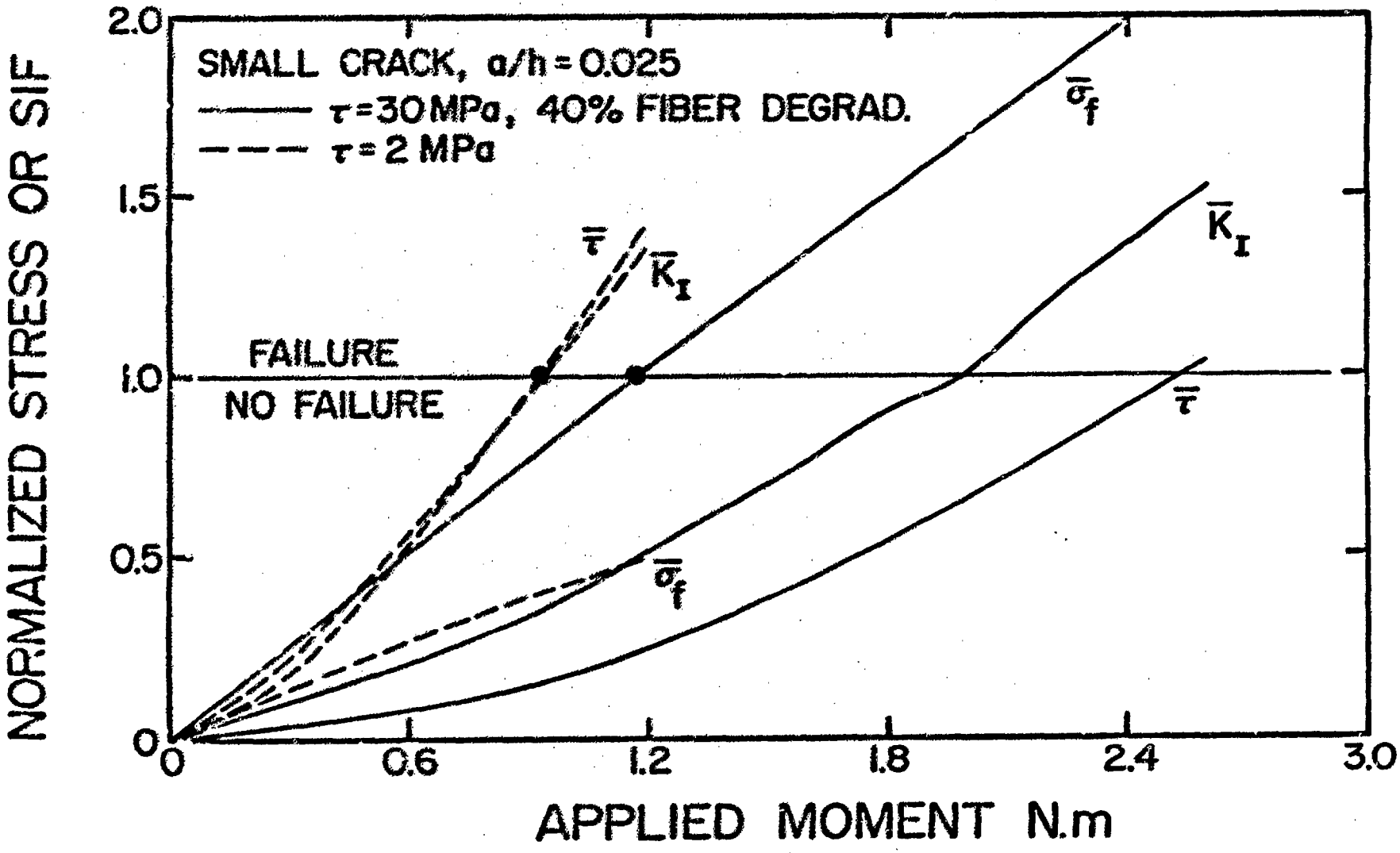


Figure 6. Failure Modes for Small Cracks.

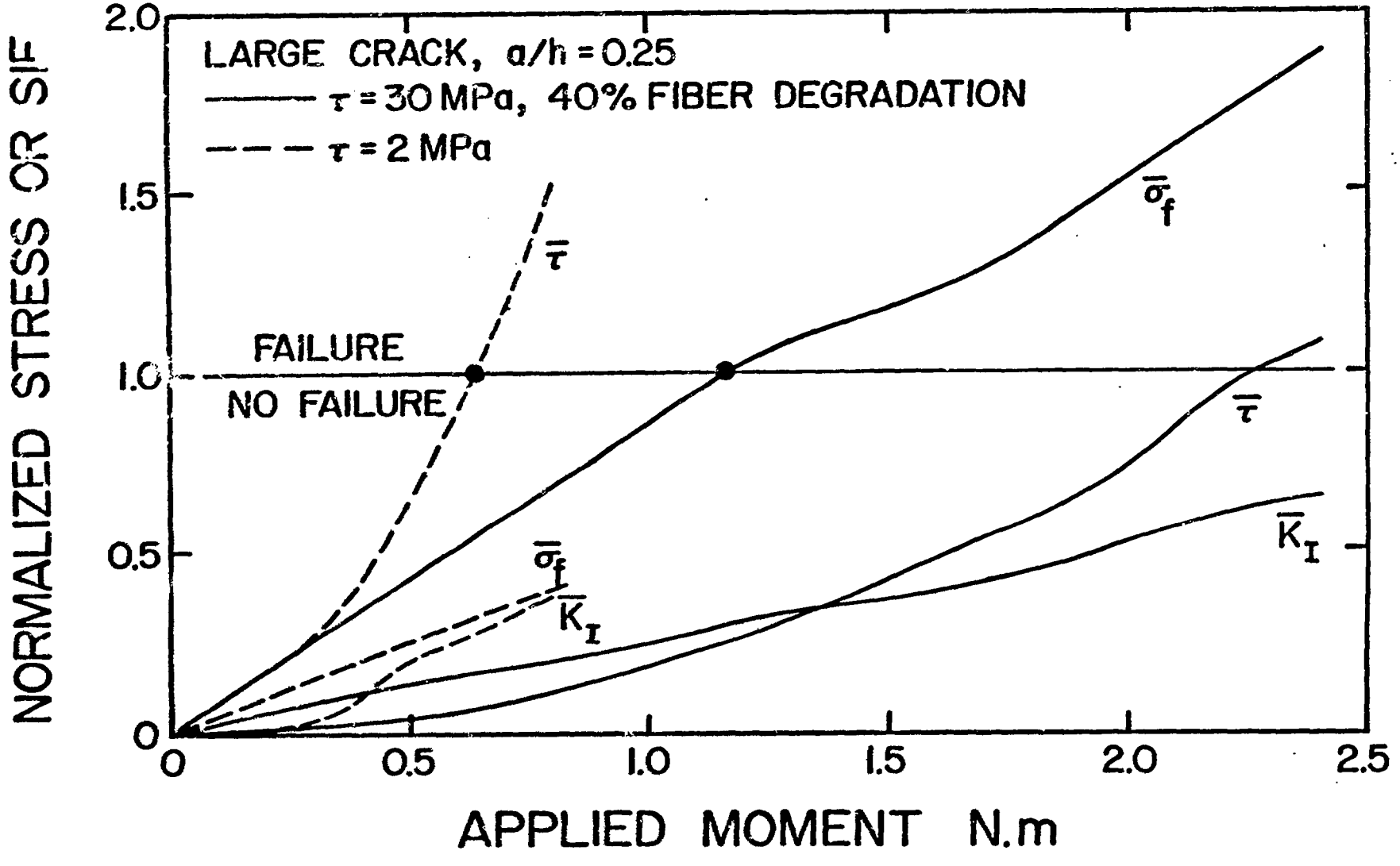


Figure 7. Failure Modes for Large Cracks.

debonding is significant. At elevated temperature, however, the failure mode would still be that of fiber fracure. A summary of failure modes as a function of interface strength, fiber strength, and crack size is presented in Table-I for the composite system discussed in this section.

Fig. 8 shows the effect of multiple matrix cracking on the compressive stress in the beam. Even though the compressive failure may never precede the three failure modes discussed earlier the results show that the stresses calculated using the flexure formula are significantly different from the true stresses.

In conclusion, the analytical model presented in this paper is capable of predicting the observed failure modes at both room and elevated temperatures by accounting for the changes in microstructural properties of the constituents and the interface condition. The predictions agree qualitatively with the limited available experimantal data [13 – 16] on brittle matrix composites.

ACKNOWLEDGEMENTS

This work was performed under the AFOSR/UES Mini Grant Program, F49620-85-C-0013/SB5851-0360. The author is grateful to Dr. Nicholas J. Pagano, Materials Laboratory, Wright-palterson Air Force Base, for his support and discussions during the course of this work.

REFERENCES

1. Wiederhorn, S.M., "Brittle Fracture and Toughening Mechanisms in Ceramics", Ann. Rev. Mat. Sci., 14, 1984, pp. 373-403.

2. Faber, K.T., "Toughening Mechanisms for Ceramics in Automotive Applications ", Ceram. Eng. Sci. Proc., 5, 1984, pp. 408-439

3. Sambell, R.A.J., D.H. Bowen, and D.C. Phillips, "Carbon Fiber Composites with Ceramic and Glass Materices", J. Mater. Sci., 7, 1972, pp. 663-675

4. Sambell, R.A.J., A. Briggs, D.C.Phillips, and D.H. Bowen, "Carbon Fibre Composites with Ceramic and Glass Matrices", J. Mater. Sci., 7, 1972, pp.676-681.

5. Phillips, D.C.,"The Fracture Energy of Carbon-Fibre Reinforced Glass", J. Mater. Sci., 7,1972, pp 1175-1191

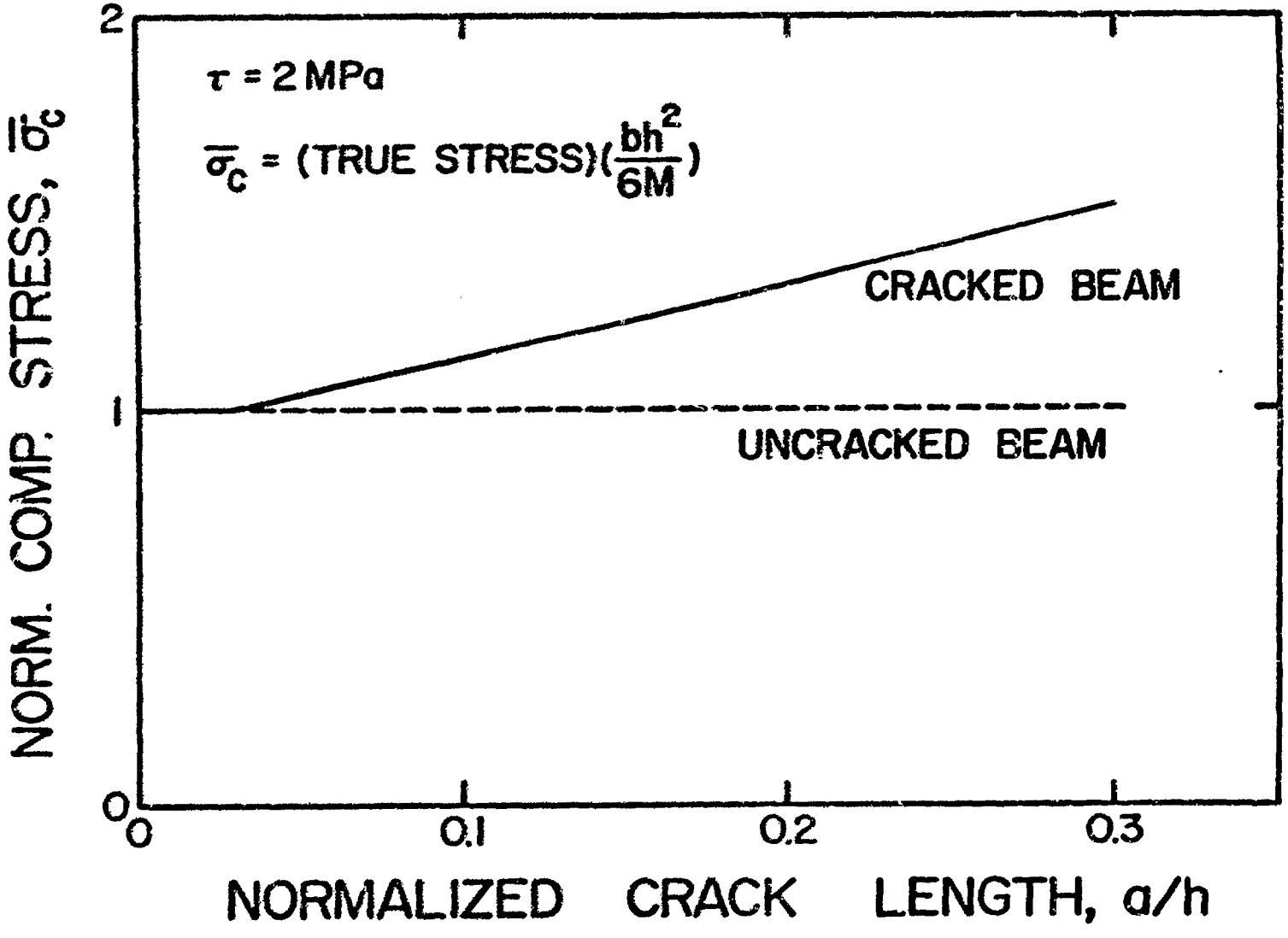


Figure 8. Effect of Multiple Matrix Cracking on Compressive Stress.

Table I. Summary of Failure Modes

Interface condition	Small crack	Large crack
Weak	Self similar cracking	Debonding
Strong with no fiber degradation	Self similar cracking	Fiber break
Strong with 40% degradation	Fiber break	Fiber break

6. Phillips, D.C., R.A.J. Sambell, and D.H. Bowen, "The Mechanical Properties of Carbon Fibre Reinforced Pyrex Glass", J. Mater. Sci., 7, 1972, pp.1454-1464

7. Phillips, D.C.,"Interfacial Bonding and the Toughness of Carbon Fibre Reinforced Glass and Glass-Ceramics", J. Mater. Sci., 9, 1974, pp.1847-1854

8. Brennan, J.J., and Prewo, K.M., "Silicon Carbide Fibre Reinforced Glass-Ceramic Matrix Composites Exhibiting High Strength and Toughness", J. Mater. Sci., 17,1982, pp.2371-2383

9. Aveston, J., C.A. Cooper, and A. Kelly, "Single and Multiple Fracture", in The Properties of Fibre Composites, Conference Proceedings, NPL, 1971, pp.16-26

10. Aveston, J., "Strength and Toughness of Fibre-Reinforced Ceramics", in The Properties of Fibre Composites, Conference Proceedings, NPL, 1971, pp.63-74

11. Aveston, J. and A. Kelly, "Theory of Multiple Fracture of Fibrous Composites", J. Mater. Sci., 8, 1973, 352-362

12. Aveston, J. and A. Kelly, "Tensile First Cracking Strain and Strength of Hybrid Composites and Laminates", Phil. Trans. R. Soc. Lond., A 294, 1980, pp.519-534

13. Marshal, D.B. and A.G. Evans, "Failure Mechanisms in Ceramic-Fiber/Ceramic-Matrix Composites", J. Am. Ceram. Soc., 68(5), 1985, pp.225-231

14. Mah. T., et al., "Room Temperature Mechanical Behavior of Fiber-Reinforced Ceramic-Matrix Composites", J. Am. Ceram. Soc., 68(1), 1985, pp. c-27 - c-30.

15. Luh, E.Y. and A.G. Evans, "High-temperature Properties of a Ceramic Matrix Composite", J. Am. Ceram. Soc., 70 (70), 1987, pp.466-469

16. Mah. T., et al., "High-Temperature Mechanical Behavior of Fiber Reinforced Glass-Ceramic-Matrix Composites", J. Am. Ceram. Soc., 68 (9), 1985, pp. c-248 - c-251

17. Marshall, D.B., B.N. Cox, and A.G. Evans, "The Mechanics of Matrix Cracking in Brittle-Matrix Fiber Composites", Acta Matall., 33(11), 1985, pp.2013-2021

18. Parker A.P., "The Mechanics of Fracture and Fatigue", E.&F.N. SPON ,1981, pp.52-62

19. Benthem J.P., W.T. Koiter in "Method of Analysis and Solutions of Crack Problems" (edited by G.C.Sih), Noordhoff International Publishing, 1973, pp. 168-170

20. Chiang, C.R., "Inhomogeneity Effect on the Stress Intensity Factor ", J. Composite Mater., 21 1987, pp. 610-618

21. Dharani L.R., W.F. Jones, and J.G. Goree, "Mathematical Modeling of Damage in Unidirectional Composites", Engr. Fract. Mech., 17(6), 1983, pp. 555-573

FINAL REPORT NUMBER 21
REQUESTED A NO-COST TIME EXTENTSION
TO BE SUBMITTED IN 1987 MINI-GRANT FINAL REPORT
Dr. Peter Disimile
760-6MG-075

FINAL REPORT NUMBER 22
REQUESTED A NO-COST TIME EXTENTSION
TO BE SUBMITTED IN 1987 MINI-GRANT FINAL REPORT
Dr. George Doyle
760-6MG-006

1986 USAF-UES RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

SENSITIVITY OF MESOSCALE WIND SYSTEMS TO VARIATIONS IN VEGETATION CANOPY PARAMETERS AND SURFACE PROPERTIES

Prepared by:	Patrick T. Gannon, Sr.
Academic Rank:	Associate Professor
Department and University:	Department of Meteorology Lyndon State College
Research Location:	Lyndon State College and Air Force Geophysics Laboratory Hanscom AFB, MA
USAF Researcher:	Dr. Robert M. Banta
Date:	January 14, 1988
Contract No:	F49620-85-C-0013/SB5851-0360
UES Project 760	S-760-6MG-100

SENSITIVITY OF MESOSCALE WIND SYSTEMS TO VARIATIONS IN VEGETATION CANOPY PARAMETERS AND SURFACE PROPERTIES

by

Patrick T. Gannon, Sr.

ABSTRACT

The two broad research areas investigated during the 1986 Summer Faculty Research Program were also addressed during the 1987 Research Initiation Program (RIP). The work at the Air Force Geophysics Lab (AFGL) concentrated on the effects of soil moisture, ridge heights and inversion strength on nocturnal katabatic flows. Simulations of these flows were conducted using a two-dimensional version of the Colorado State University (CSU) Cloud Mesoscale Model (26).

The specific focus of this research was on the effects of soil type and soil moisture on surface temperature, heat flux and downslope wind speed.

The second research area was the area of major concentration in terms of resources and time. The core task was the integration of a vegetation canopy parameterization into a mesoscale numerical model used primarily to simulate sea and lake breezes.

The AFGL study proceeded well and conclusions were drawn on a timely basis. The canopy effect was adversely affected by

numerical problems associated with incompatibilities between model predictive and canopy iterative schemes. Several of the intended experiments could not be conducted as planned.

I. INTRODUCTION

The research contract between Lyndon State College and Universal Energy Systems (UES) stipulated the period of performance as 15 January 1987 through 31 December 1987 with the Final Report due 31 December 1987 and dated 31 January 1988. On 14 October 1987, a request for an extension of the reporting date to 14 January 1988 was made to UES.

This report provides a chronology of administrative and research activities, important contacts made and milestones reached. The report provides a discussion of the methods used and the conclusions derived from the research in the two study areas. The chronology reflects the overlapping of efforts in the two research areas.

II. CHRONOLOGY

15 Dec 86	Contacted Dr. Ir. J. Goudriaan at Landbouw-universiteit, Wageningen, The Netherlands concerning the availability of a FORTRAN version of his crop micrometeorology program.
15 Dec 86 - 20 Feb 87	Diagnosed numerical and graphical output from CRAY model runs. The focus of this effort was to describe the dynamics of the predicted katabatic flow fields resulting from early sensitivity experiments. Draft reports were produced describing the magnitude of the terms in the

governing equations, the locations along the mountain slope where their relative importance is maximized and the overall linkage to the initial ridge height, soil type, temperature and moisture, and initial ambient atmospheric conditions. Potential temperature, temperature, vertical motion and perturbation downslope velocity patterns and values were studied and tentative conclusions drawn.

Several factors converged such that this investigator reduced his efforts in this area and redirected them in others. It was decided at AFGL that detailed analysis of the output would be apportioned such that I would concentrate on soil type and moisture effects while Dr. Banta and Capt. Kim C. Pantley (also of AFGL) would concentrate on remaining areas. This arrangement facilitated the preparation of papers for the Mountain Meteorology Conference scheduled for August 87. The impending research effort with the canopy program was also a factor. This decision was fortitious as it was later discovered that magnitudes of some output variables were spurious, even though the overall patterning was realistic. Plans for new model runs at AFGL were not implemented until Fall 87 when funds were released for additional CRAY model runs. These runs were for

an expanded two-dimensional domain. As of the date of this Final Report, plans call for my contribution to a technical report on the new findings which will be prepared after the termination of this contract.

14 Jan 87 — Received from Dr. Ir. J. Goudriaan, Landbouwuniversiteit, Wageningen, The Netherlands the BASIC vegetation canopy program developed for a personal computer by Dr. Chen Jialin.

15 Jan 87 — Secured the services of senior student Peter Clement to perform computer program editing and related functions.

15 Jan 87 - 30 Jun 87 — Coordinated frequently with Ms. Abigail M. Thomas, Lyndon State College Coordinator of Academic Computing, and Sgt. Charles Crouch, Computer Systems Specialist, AFGL. Purpose of this interaction was to obtain in VAX-compatible format FORTRAN and data files which were in CYBER-generated unformatted code. Some of these files were on disc and tape at AFGL; others were on tapes which were generated in 1981 at the Environmental Research Laboratories in Boulder, CO. The files and the services of Ms. Thomas and Sgt. Crouch were absolutely essential to the canopy project.

20 Jan 87 - 10 May 87	Supervised the conversion of the Jialin BASIC version of the canopy program into a VAX-compatible FORTRAN version which would mesh with the sea breeze model. Peter Clement performed this function with guidance from Ms. Thomas. The PC version was run successfully.
11-14 Feb 87	Made the first of two trips to AFGL, Hanscom AFB, MA to confer with Dr. Robert Banta on my role in the downslope flow numerical modeling experiments.
1 Mar 87 - 15 May 87	Microfiche records of numerical output generated by CRAY computer were examined using the college microfiche reader. From these records temperature, moisture, and radiation values were extracted for graphical depictions of surface influences on katabatic flows.
10 May 87 - 17 Jun 87	Tested the VAX program frequently and encountered unexplained overflows. This problem slowed the total effort down while line-by-line searches of the code did not reveal the problem. Ms. Thomas finally identified the problem as one associated with differences in DO LOOP incrementing which exist between PC's and the VAX. Finally a VAX run was produced which replicated the PC output.
26 May 87 - 30 May 87	Made the second of two trips to AFGL, Hanscom AFB, MA to confer with Dr. Banta. Primary purpose of trip was to collaborate on preparing a paper to be delivered at the American Meteorology Society

Mountain Meteorology Conference scheduled for 25-28 August 1987. During this visit, I prepared drafts of certain portions of the enclosed preprint titled Soil Moisture and Ridge Height Effects on Katabatic Flow Simulations (atch 1). My principal contributions were the graphical displays in Fig. 5, Fig. 7, para 5 and para 7. These contributions were based upon the activities mentioned during the period 1 March - 15 May 1987.

1 Jun 87 - 31 Dec 87: Integrated the canopy program into the sea breeze model. Modified canopy program to function as a subroutine. Adjusted code and physical units. Debugged entire package. Ran several experiments.

15 Dec 87 - 12 Jan 88: Prepared Final Report for UES.

III. OBJECTIVES OF THE DOWNSLOPE FLOW EFFORT

The primary objective of this effort was to investigate the response of downslope flows to several soil types and soil moisture distributions.

A related objective was to analyze numerical output from CRAY model runs of the CSU model to determine linkages among motion fields and soil parameters.

These objectives were met and conclusions were drawn from the model runs.

Another objective was to incorporate the canopy algorithms into the CSU model and run experiments from nocturnal katabatic and upslope convective episodes. This objective was not attained as the canopy project was complex and beset with setbacks. Also, funds were not available to AFGL to enable this aspect of the project to be completed.

IV. DOWNSLOPE FLOWS: METHODS AND FINDINGS

The methods and findings are described in Banta et al (4). A summary is presented here to highlight this investigator's contribution.

DOMAIN: Hill; 750m (ht) x 14km (wide); symmetric

RESOLUTION: 100m (horizontal); 0.5k Pa (vertical)

EXPERIMENT A: Soil Moisture Comparisons (Hi vs Lo)

a) soil type - silty clay loam

b) location: on slope near base at x = 4km

c) Hi soil moisture: approx. 0.48 cm^3 cm^{-3}

d) Lo soil moisture: approx. 0.20 cm^3 cm^{-3}

Fig. 5 (atch 1) shows the rapid drop in surface temperature (T_{sfc}) in the case of Lo soil moisture compared to the Hi soil moisture simulation. Throughout the 6-hr simulation period, the Lo soil moisture temperature was about 4k colder. The downward heat flux ($w'\theta'$) was also higher for the Lo case indicating a more pronounced surface inversion than for Hi. The dynamic significance of the drier soil is reflected in the higher value of peak u for the Lo case.

At the five hour point, changeovers occurred in Lo $w'\theta'$ and peak u. These can be attributed to the developing cold pool at the base of the hill. This cold layer is air which moved downslope during the night and was reinforced by further radiational cooling. It produced a reverse upslope pressure gradient which caused a convergence zone near x = 4 km. This weakened the inversion and consequently downward heat flux. Peak u for Lo declined as the convergence zone moved upslope (see Fig. 4, atch 1).

EXPERIMENT B: Soil Moisture/Type Comparisons

a) Soil Types: silty clay loam and sand

b) Location: on slope near base at x = 4 km.

c) Hi Soil Moisture:

1. silty clay loam: approx. 0.48 $cm^3\ cm^{-3}$
2. sand: approx. 0.4 $cm^3 cm^{-3}$

d) Lo Soil Moisture

1. silty clay loam: approx. 0.12 $cm^3\ cm^{-3}$
2. sand: approx. 0.1 $cm^3 cm^{-3}$

The table on p. 143 (atch 1) describes initial soil thermal and moisture properties for the two soils.

For both soils, Hi soil moisture values resulted in reduced surface cooling rates compared to the Lo case. The higher values of hydraulic conductivity (K_n), water diffusivity (D_n) and thermal diffusivity (λ) in the case of sand resulted in a more rapid upward diffusion of moisture and heat then was observed for silty clay loam. The loss of heat to space for the sand soil

layers was more abrupt (Fig. 7, atch 1) compared to silty clay loam (fig. 5, atch 1).

The most rapid surface cooling was observed for Lo soil moisture and sand. The course-grained sand has a higher λ for about the same soil moisture content (.099 cm^3cm^{-3}) as silty clay loam (.119 cm^3cm^{-3}).

A major factor which acts to control the rate of surface cooling is downward long wave radiation. The moisture gained by the surface layer from surface evaporation is higher in the case of silty clay loam since the soil moisture profile is higher. A separate diagnostic study (not replicated here) verified that absorption and re-emission of long wave radiation was a significant factor in reduction of the ground surface cooling rate for silty clay loam.

Figure 7, atch 1 shows the most rapid cooling occurred during the first hour in the case of Lo soil moisture for sand. The T_{sfc} profile shows an extreme surface temperature drop of almost 15C in six hours. The downward heat flux ($w'\theta'$) reaches a maximum value at three hours. The reversal of this trend is dynamically induced as in the case of Lo soil moisture for silty clay loam. The most vigorous peak u and backflow were seen for the Lo sand case.

V. OBJECTIVES OF THE VEGETATION CANOPY EFFORT

The primary stated objective was to incorporate vegetation canopy algorithms into the CSU Cloud/Mesoscale Model (26). The

code was to be tested with vegetation parameters appropriate for a Rocky Mountain or Florida environment.

The original intent was to test the canopy parameterization developed by Deardorff (6) and employed by McCumber (19). Original plans called for a modification of the iterative scheme used in (19) so as to replace it with a predictive scheme developed by Tremback and Kessler (24) for the surface energy budget. Research into the vegetation problem focused on the simulation study of crop micrometeorology by Goudriaan (11). The complexities of the energy flows and turbulent transfer processes were numerically simulated by a CSMP language model. This investigator inquired of Dr. Goudriaan as to the availability of a FORTRAN version. He indicated that a version existed which had been developed by Japanese crop micrometeorologists (12). He kindly provided me with a version written in BASIC which could be run on a PC. This version was developed by Chen Jialin and it together with documentation is included as atch 2. The matrix inversion technique together with other physical improvements resulted in a PC program which had a time step of 450 seconds and a total execution time of 30 minutes to simulate a day (at one location). Certain simplifying assumptions were made by Jialin which enabled this large time step model to replicate results from the CSMP and FORTRAN versions having a time step of about one second (p. 58, atch 2).

The PC version was adopted for modification as a vegetation canopy subroutine in the 2-D sea breeze model of McCumber and Pielke (20) as modified by Gannon (8). Motivation for this program adoption was provided by:

1. Recognition of the computer economies which the Jialin program offered.
2. Recognition after initial research that it was unlikely a scheme could be developed from the Deardorff version which would be cost effective and offering the same physical completeness as the Jialin version.
3. Recognition of the multiple crop and vegetation canopy properties which could be input into the Jialin version.
4. Recognition that it was unlikely mesoscale numerical modellers would wish to incorporate canopy algorithms into existing models if time steps were on the order of a few seconds.

VI. IMPORTANT FEATURES OF THE JIALIN VERSION

1. Direct, diffuse, and long wave radiation formulations.
2. Vector-matrix representation of radiation transfer through canopy.
3. Large time step.
4. Energy flows treated like electrical resistances.
5. Uncoupled multi-layer model-use of enthalpy flux density and saturation heat flux density.
6. Soil layer-temperatures iteratively determined with linkage to air temperature within canopy.

VII. MAJOR MODIFICATIONS TO THE JIALIN VERSION

1. Compatibility of units: As the sea breeze model is in cgs units and the canopy program is in mks units, extensive unit conversions among shared variables were necessary. These conversions proved to be occasional sources of error. Production runs were delayed as a result.

2. Crop height (Z_c) and reference height (Z_r): In both the Goudriaan and Jialin versions, the input atmospheric, soil, and radiation data were observationally-based. Z_c was 2.5m (maize) and Z_r set at 3m. In theory, it was shown by Goudriaan that increasing the height of the reference level up to 5m resulted in minor changes in output, but no instabilities. In the present study, all input data to the canopy program originated from the 2-D sea breeze model. It was deemed necessary to test the canopy algorithms using the original canopy parameters and heights, thus subjecting the entire package to severe test conditions. The challenge came about from the staggered level configuration in the sea breeze model (atch 3). Wind velocity (u) is defined at 17.5m and nine higher levels, pressure (P) is defined at the surface, 17.5m and the higher levels. Temperature (T) and mixing ratio (Q) are defined at the surface, top of the laminar sub-layer (about 4cm), 25m and higher levels. Initial attempts to specify Z_r for u at 17.5m

and T, Q at 25m resulted in unrealistic resistance values between Z_c and Z_r and the canopy/soil layer iteration scheme failed to converge.

After a number of interpolation schemes were tried, a final solution consisted of computing u, T, and Q at Z_r = 3m as a function of u_* (friction velocity), T_* (friction temperature) and Q_* (friction mixing ratio) and appropriate profile functions. The formulations were those used by McCumber (19). When Z_c and Z_r were set to 17.0m and 17.5m (trees), the model performed well without modification.

The formulations are identified in atch 3. Atch 3 is a copy of the complete 2-D sea breeze model including the vegetation algorithms. Comments are interspersed at appropriate locations to inform the reviewer of the functions that segments of code perform.

3. Soil moisture and transpiration. The original canopy program treated the soil as always being at field capacity (soil water potential = -0.1 bar). The values for soil thermal conductivity and volumetric specific heat were compatible. Two of the experiments reported on in this paper used the same values.

For these two experiments, the canopy soil temperature and moisture regimes were treated as totally separate domains from water, bare soil and urban areas. The complete soil temperature and moisture algorithms were applied to the latter two surface types; however, it was not considered cost and time effective to

cycle through these algorithms a third time for canopy soil layers. This should be done in a future research project.

For a third experiment, the grid square bare soil layer values of soil moisture, soil temperature and soil volumetric specific heat were applied to the canopy soil layers. Many of these grid square soil layer domains were not at saturation. In these runs, the soil water potential is computed as the average of this property over the top seven soil layers.

4. Resistance to evaporation: the original canopy program specified a resistance to evaporation in the in-canopy layer adjacent to the surface as zero sm^{-1} (compatible with a continuously wet soil). This procedure was followed for the first two model runs which treated the soil as saturated. For the third run where soil layers were not at saturation, this resistance had to be modified. Thus, the resistance had to be altered as some function of soil moisture content in the top soil layer. The only available guidance for a resistance value for a very dry soil was that found in Goudriaan (11). An appropriate value was $10^6 sm^{-1}$. Interpolation was accomplished using the interpolation subroutine of Jialin (15) and the above-mentioned resistance value. This proved to be too large and a smaller value was adopted.

The independent variable was specified as η_s/η (see atch (1)). When $\eta = \eta_s$ (saturation) the resistance is zero. When $\eta = \eta_{wilt}$ (wilting point) the resistance was set at 1000. The interpolation routine fixed the appropriate value between these limits.

5. Radiation: In the original Jialin version, radiation could be specified as net (short and long wave) or global (short wave) and long wave separately. As the sea breeze model computes short and long wave radiation, these values are passed to the canopy subroutines. The short wave radiation is processed by the canopy program into direct and diffuse components. Some uncertainties still exist in accounting for the 'fraction of overcast sky' effect described by Goudriaan and Jialin. Further work needs to be done here.

6. Special variables: Many DATA statements in the original canopy program specify pairs of independent and dependent variables which SUBROUTINE INTERPOL operates on to define a current value. For example, the relationship defining leaf temperature (T2) and maximum net CO_2 assimilation rate (F2) has the maximum possible T2 value as 40C. Since the canopy program was not developed for a tropical environment (e.g., south Florida), special coding had to be inserted to prevent numerical problems when maximum values were exceeded.

VIII. FLOW SEQUENCE, CANOPY PROGRAM

In lieu of a flow chart, a series of statements is provided which describe the main computational procedure for the vegetation canopy algorithms.

SEA BREEZE MODEL (Subr. SFCHT) CALLS VEGGE

VEGGE:	computes	-	time step
		-	vapor pressure
		-	soil layers, conductivity, specific heat
		-	reads input parameters
		-	soil temperature
		-	radiation, turbulence parameters
	CALLS	-	DIFFRAD (extinction and reflection coefficients of diffuse radiation)
	computes	-	initial values of heat flux density into soil, canopy relative water content, actual canopy water content
	redefines	-	wind, temperature, vapor pressure from sea breeze model
	computes	-	variables used in turbulence subroutine (AERODYNAM)
	CALLS	-	AERODYNAM (turbulent resistances and in-canopy winds)
	computes	-	net long wave radiation at canopy top
	computes	-	sun height dependent parameters, fraction of overcast sky, magnitudes of direct and diffuse visible and near-infrared radiation
	computes	-	leaf absorption, leaf resistance

computes - enthalpy flux source

ITERATION - (heat flux density into soil surface)

- CALLS FLUXES: Monin-Obukhov length, saturation heat source, in-canopy air temperatures, soil temperature in top layer.

CALLS - DEWSOILTMP (leaf wetness, dew, soil water potential, canopy water content and soil layer temperatures).

PRINT - output data

compute - Canopy top θ_*, U_*, Q_* (canopy contribution to sea breeze model for computation of grid square surface layer fluxes).

CALLS - PROFILE (data output every 3 hours; computes leaf temperatures, vapor pressure profile and vapor pressure deficit profile).

END

IX. SEA BREEZE MODEL: BACKGROUND

The vegetation canopy algorithms were incorporated into the 2-D sea breeze model to test the sensitivity of mesoscale circulations to vegetation as well as bare land, water, and urban surfaces. Soil types and soil moisture distributions also influence the location and intensity of convergence zones associated with variable surfaces. McCumber (19), Gannon (8) and others have demonstrated that surface properties need to be

specified with care for accurate predictions of 'dry' mesoscale systems and 'moist' meso-convective systems.

The objective of incorporating the canopy algorithms into the CSU model could not be met. Time and funding constraints precluded the attainment of this objective. However, the testing of the combined sea breeze-canopy model on the Lyndon State VAX 11/750 proceeded satisfactorily with most of the major technical interfacing problems resolved.

X. SEA BREEZE MODEL: ELEMENTS

The model (without vegetation) is very similar to that described in the Technical Report (8) which is included as atch 4. Modifications include:

1. Horizontal and vertical diffusion coding used by McCumber (19).
2. Soil moisture and temperature prediction coding developed by Tremback and Kessler (25) and used in the CSU model (26).

Principal elements of the complete model are:

1. MAIN - all of the dynamic formulations which predict horizontal and vertical velocities, pressures, potential temperatures, and moisture.
2. OZONE - computes attenuation of direct solar radiation through the ozone layer.

3. COOL - computes upward and downward IR fluxes at all model levels.

4. SOLRAD - computes probabilities of solar radiation passing through cloud fields depending upon solar zenith angle.
   - computes effective albedo of the surface as a function of solar zenith angle.

5. FIRCAT, SECCAT, THRCAT - computes transfer of direct and diffuse solar radiation as a function of spectral interval, absorbing gas, cloud field and solar zenith angle.

6. SFCHT - redefines soil moisture and temperature for next prediction cycle.
   - calls SOILW1 (moisture values for soil thermal properties).
   - calls SOILT (predicts new soil layer temperature and urban cultural feature layer temperature).
   - calls SOILW2 (predicts new soil moisture values).
   - calls VEGGE (predicts new canopy properties to include variables used to update fluxes in MAIN).

XI. SEA BREEZE MODEL: WEIGHTING FACTORS

Unlike the CSU model, where horizontal resolution is on the order of 100m, the sea breeze model has a grid mesh of 11km. In the CSU model, a simple allocation of a single physiographic feature (e.g., trees, urban, bare soil, etc.) to a grid square is possible. The dynamic equations as well as the surface energy budget formulations would be much simpler. In the present case,

each grid square can contain several physiographic features. This requires weighting factors for water, land, urban, bare soil, and vegetated regions. These factors are applied in such a way that grid square values of temperature, wind, and moisture reflect the weighted contribution of each type of surface. Atch 5 lists the weighting factors and soil types for each grid square.

XII. SEA BREEZE MODEL; ATMOSPHERIC AND SOIL PROFILES

The synoptically undisturbed case study day chosen for the testing of the vegetation canopy subroutine was July 16, 1975. The initial sounding profile for the model runs is shown in atch 6. The initial soil temperature profiles for bare soil, urban and canopy are shown in atch 7. Soil moisture values were computed from rainfall, wind, and evaporation data. They are listed as a DATA statement in atch 3.

XIII. SEA BREEZE MODEL: DOMAIN

Atch 8 is a schematic of the horizontal domain of the three-dimensional version of the sea breeze model. The 2-D slice used in the present investigation is for the array labeled 22 (ordinate). Here the first grid point containing land is at abscissa point 6 and the last is at 26. Surface properties for this slice are considered to be quite realistic; especially the soil moisture, soil type, albedo, and urban coverage. Vegetation cover is realistic. Vegetation types are altered for sensitivity studies.

XIV. COMPUTER MEMORY REQUIREMENTS

The combined sea breeze and vegetation canopy program consists of a main component and twenty-two subroutines. Text and data require 168,000 decimal words. Data requirements amount to 78,000 words. This is on the high side for a two-dimensional model, yet is understandable when radiation, soil layer, and vegetation parameters and variables are being used.

XV. RESULTS OF NUMERICAL EXPERIMENTS

Early experiments for Florida conditions were encouraging. The basic model time step was 90 sec., the radiation time step 900 sec. and the canopy time step 450 sec. Emphasis was then shifted to daytime runs to test the impact of vigorous radiation, thermal and moisture fluxes. From an investigative standpoint, these were more interesting in terms of mesoscale circulations. Problems were encountered in many cases one hour (real time) after sunrise. The iteration scheme failed to converge and on some occasions the predicted air or soil temperatures were far too unrealistic and oscillated badly. The multiplicity of variables in the canopy iterative scheme and the feedbacks into the main model algorithms made error diagnosis difficult and time-consuming. Near the end of the RIP research period, the main sources of difficulty were identified and remedial action taken. Problem areas which were identified and for which solutions have been found are:

1. Defining the reference level for short vegetation. This previously discussed problem was resolved by using profile functions in the surface layer.

2. Rapidly changing values of thermal conductivity and specific heat for unsaturated soils. Assuming soils to be saturated was a partial solution.

3. Time step. With a tropical daytime radiation field, a time step of 450 sec. simply will not work. Reduction of the time step to 90 sec. for a saturated soil and 22.5 sec. for unsaturated soils represented partial solutions.

4. Radiation field. To achieve stable computations and realistic output, an average low cloud cover of 0.6 across the peninsula was prescribed. For the day in question, this was not unrealistic as cloud cover ranged from 0.3 to 1.0 during the day. The reduction in incoming radiation also slowed down the adverse effects of other variables mentioned.

5. Resistance to evaporation. The only available values were those published by Goudriaan (11). A formulation relating soil moisture or soil moisture potential to resistance could not be found. For experiment 3, the values of resistance had to be reduced three orders of magnitude and coupled with a time step of 22.5 sec. before a stable run could be obtained.

A. EXPERIMENTAL DESIGN

Four model experiments were run. Experiment 1 simulated sea and lake breeze mesoscale circulations where the canopy cover (atch 5) was similar to sawgrass with a crop height of 2.5m.

Experiment 2 was identical to experiment 1 except that the

canopy cover is similar to Florida pine trees with a high leaf area index appropriate to thick palmetto stands within the canopy. The tree tops are at 17m with a reference height of 17.5m.

Experiment 3 was identical to experiment 2 except that the grid square soil moisture and associated thermal properties were also applied to the sub-canopy soil layer. This experiment will simulate for unsaturated grid squares varying degrees of 'foliage fever' (6). Even though the root drawdown algorithm is invoked, the soil moisture fluxes are computed using the bare soil formulations in SOILW1, SOILW2, and SOILT. Therefore, the experiment is a hybrid experiment.

Experiment 4 is a run where the vegetation canopy is removed. This run is used to show how the mesoscale circulations evolve in the absence of vegetation. The direct and subtle effects of a canopy can thus be diagnosed.

The time chosen for display of results is 1600 LST. At this time the double Florida Sea Breeze is well-established with the Gulf Coast circulation being the strongest. The canopy influence can only be assessed by intercomparison of the four model run results. In the actual Florida afternoon situation, concentrated lines of thunderstorms with downdrafts and differential cloud shielding would alter the predicted convergence zones simulated by the model. The focus of the reader should be on the sensitivity to changes resulting from variable model input conditions.

Although the figures presented show only that portion of the vertical domain extending up to 1.527km (about 850mb), motion and

scalar perturbations extend much higher. The computer printouts for the experiments are enclosed as atch 9. These can be used by the reader for additional insight into model results. Special output of canopy variables is included in these computer printouts.

B. INITIAL CONDITIONS: July 16, 1975

Sunrise: 0547 EST

Cloud cover: 6/10

Surface temperature: see Fig. 1

Potential temperature: see Fig 2 and atch 6

Moisture: see Fig 3 and atch 6

Wind: see Fig. 4 and atch 6

Grid Square soil moisture: see atch 3

C. EXPERIMENT 1

Model time step: 90 sec.

Radiation time step: 450 sec.

Canopy time step: 90 sec.

Canopy soil moisture: -0.1 BAR

Canopy soil thermal conductivity: 1.3 $Jm^{-1}s^{-1}K^{-1}$

Canopy soil specific heat: 2×10^6 $Jm^{-3}K^{-1}$

Foliage height: 2.5m

Reference height: 3.0m

Figs. 5-9 show predicted fields for surface temperature, potential temperature, moisture, wind, and vertical velocity.

1. Surface temperature Fig. 5 shows maximum heating on both coasts. It is difficult to separate the urban from the canopy effects on the Gulf Coast as the

convergence zone here contributes to heat and moisture concentrations. The urban influence can be seen on the Atlantic coast.

2. Potential temperature. Fig. 6 shows the principal unstable potential temperature profiles at grid points 6-10 and 21-24. Warming is seen across most of the domain with maxima on the Gulf Coast and east of the lake.

3. Moisture. Fig. 7 shows considerable moistening across most of the domain with elevated values reaching well beyond 1.527km. Again, the highest concentrations are vertically oriented over the Gulf Coast and the weaker convergence zone east of the lake. An important question to be answered is: How much of this moisture is due to evapotranspiration and how much to advection of moisture from the Gulf, Atlantic, and inland water bodies? Experiment 4 was run to assess the transpiration contribution.

4. Wind. Fig. 8 shows well-organized convergence zones on the Gulf Coast and the eastern shore of the lake. Wind speeds have doubled in the convergence zones from the surface to about 1km. Weaker winds are seen in the central part of the domain. These are associated with sinking motion in the center of the peninsula.

5. Vertical Velocity. Fig. 9 shows well-organized vertical motion maxima (upward) from 1.072 km to 1.527 Km with weaker minima (downward) on the periphery. As the initial vertical motion field is at

rest, this figure shows the significant role of the surface in establishing preferred regions for convection.

D. EXPERIMENT 2

Model time step: 90 sec.

Radiation time step: 450 sec.

Canopy time step: 90 sec.

Canopy soil moisture: -0.1 BAR

Canopy soil thermal conductivity: 1.3 $Jm^{-1}s^{-1}K^{-1}$

Canopy soil specific heat: 2×10^6 $Jm^{-3}K^{-1}$

Foliage height: 17.0m

Reference height: 17.5m

Figs. 10-14 show predicted fields for surface temperature, potential temperature, moisture, wind, and vertical velocity.

1. Surface temperature. Fig. 10 shows maxima in the same locations as Experiment 1, but less in magnitude.
2. Potential temperature. Fig. 11 shows a warmer core above the Gulf Coast compared to Experiment 1. The vertical profile in the lower layers here is more adiabatic than superadiabatic as was seen in Fig. 6. It is hypothesized that this result originates from the canopy providing an elevated heat source, particularly from grid point 6 to grid point 13.
3. Moisture. Fig. 12 shows the Gulf Coast vertical moisture core retarded in its westward progression by 22 to 33 km. Surprisingly, the values of mixing ratio are slightly lower in the western portion of the domain

than was the case in Experiment 1. It is possible diffusion of the moisture from the elevated transpiration source was more effective in this case.

4. Wind. Fig. 13 shows the Gulf Coast wind maximum of $6ms^{-1}$ is both elevated and retarded compared to Experiment 1. Elsewhere, there are no major differences. Apparently, the higher vegetation produces a more effective vertical turbulent momentum transport than the sawgrass.

5. Vertical Velocity. Fig. 14 shows slightly stronger upward and downward motions over the Gulf Coast. Most striking is the retardation of the vertical motion maximum on the west coast and its inclination to the east with height. The vertical motion maximum over the lake region is not quite as well-defined as in Experiment 1.

E. EXPERIMENT 3.

Model time step: 90 sec

Radiation time step: 450 sec

Canopy time step: 22.5 sec

Canopy soil moisture: same as bare soil; changes each model time step

Canopy soil thermal conductivity: varies each model time step

Canopy soil specific heat: varies each model time step

Foliage height: 17.0m

Reference height: 17.5m

The experimental design was not as demanding as originally

intended. The maximum resistance to evaporation was specified as 1000 sm^{-1} in lieu of $10^6 sm^{-1}$. This compromise was made to insure computational stability. However, the results do indicate a more vigorous circulation resulting from higher canopy top sensible heat fluxes.

Fig. 15 shows surface temperatures about 1°C higher across the domain compared to Fig. 10.

Fig. 16 shows potential temperatures higher across the domain compared with Fig. 11. The Gulf Coast sea breeze has a low level warm core 2°C higher than in Experiment 2. The warm core is better developed and reaches higher levels. The low level atmosphere is strongly superadiabatic compared with Experiment 2. The enhanced canopy top temperatures here are representative of the 'foliage fever' mentioned earlier.

Fig. 17 shows moisture values only slightly less across most of the domain when compared to Fig. 12. The most prounounced change in moisture content is in the concentrated vertical core on the Gulf Coast. This maxima results from the vigorous sea breeze circulation generated by greater sensible heat input from the canopy despite reductions in transpiration due to unsaturated soil layers.

Fig. 18 shows maximum easterly winds at 0.617 km greater than 7 ms^{-1} on the Gulf Coast and at 0.346 km on the Atlantic coast. The convergence zones are much more pronounced; especially on the Gulf Coast. From this one experiment, it would appear that a forecaster would be justified in predicting intense localized moist convective activity on the west coast despite an

apparent overall decrease in transpiraton.

Fig. 19 shows a very strong core of upward vertical velocity on the Gulf Coast at 1.527 km (the maximum value was 57 cms^{-1}). Unlike Experiment 2, the core was vertically oriented. A possible explanation is the lag of the Gulf Coast wind maximum (u) at 1.527 km in Experiment 2 with respect to this level in Experiment 3. With vertical motions as strong as shown in Fig. 19, moisture and heat would be transported upwards to higher levels than would occur with Experiment 2.

F. EXPERIMENT 4.

Model time step: 90 sec

Radiation time step: 450 sec

Canopy: algorithms not called

Soil moisture: varies each model time step

Soil thermal parameters: varies each model time step

For this model run, vegetation parameters were removed from the code. Dynamic fields were generated with water, bare soil and urban surfaces influencing their development. Comparisons are best made by reference to Experiment 3. However, it should be remembered that the resistance in that run for the canopy layer adjacent to the soil was not allowed to exceed 1000 sm^{-1}. Thus, some caution is in order.

Fig. 20 shows surface temperatures were the highest of all the four runs. Maximum values were noted on the Gulf Coast, west of the lake and on the Atlantic coast. Variable grid square soil moisture soil type, urban coverage and dynamic feedback contributed to these values. Again, it is stressed that in the actual and the usual case, moist convective influences would have

modulated these values.

Fig. 21 shows the most vigorous 'heat island' of the four experiments on the Gulf Coast. The sensible heat acquired from the heated surfaces by the air flow is concentrated in this region.

Fig. 22 is dramatic proof that incoming radiation was apportioned into sensible heat instead of latent heat in this experiment. Many of the soil layers experienced drying and thus resulted in reduced evaporation. The moisture values are well below those of the other three experiments. This illustrates the importance of evapotranspiration in the peninsula-scale moisture supply which would be available for moist convection. Capping of the surface with cultural features and removal of vegetation will reduce the moisture available for mesoscale-generated thunderstorms.

Fig. 23 shows wind distributions different from and weaker than those resulting from canopy experiments 1-3. An exception is the easterly wind maximum centered at about .075 km east of the lake. Convergence zones are weaker than those seen in Experiment 3.

Fig. 24 shows a Gulf Coast vertical motion maximum with an upward motion core about 10 cms^{-1} less than seen with Experiment 3. The major difference from the other three experiments is the vigorous maximum on the east shore of the lake. The surface temperature differential between the lake and land surfaces is responsible for this strong vertical motion region.

XVI. CONCLUSIONS

A. SEA BREEZE CANOPY MODEL

1. The Goudriaan-Jialin canopy model is considered to be useful, physically realistic, adaptable and efficient either as a 'stand-alone' program or as a subroutine package in a mesoscale numerical model.
2. Maximum time efficiency results wtih slowly-changing moderate (e.g. nocturnal) radiation fields when the soil layer is near-saturation and when near-soil resistance is not far from zero.
3. Canopy effects are significant and should be recognized when simulating katabatic or other mesoscale (e.g., sea breeze) flows.
4. To properly specify reference heights for short vegetation, much higher vertical resolution is required in numerical models. The canopy motion and scalar fields are very sensitive to the predicted reference height values of wind, temperature and mixing ratio.

B. KATABATIC FLOW EXPERIMENTS.

1. Katabatic flows are very sensitive to soil type and soil moisture content. The most rapid surface cooling rate was observed for a dry sand layer. The lowest surface cooling rate was observed for a moist silty clay loam.
2. The higher soil moisture contents with the saturated silty clay loam soil provided an

evaporative moisture source for the lowest atmospheric layer. This had the effect of retarding land surface radiational cooling to space. The absorption and re-emission of radiation back to the surface resulted in a reduced downward sensible heat flux to the surface compared to the sandy soil. This was a reflection of the reduced surface layer inversion strength in the case of silty clay loam.

3. Downward katabatic flows and subsequent upslope flows at the base of the hill were more intense for those cases where land surface cooling was the greatest.

XVII. RECOMMENDATIONS

1. Incorporate the Goudriaan-Jialin algorithms into a model simulating moist convective processes. The CSU model would be an ideal test vehicle. Even if the canopy time step was reduced to a few seconds, the results would be valuable enough to justify the additional computer costs (e.g., CRAY computer).
2. Simulate downslope katabatic flows using the canopy algorithms. Most complex terrain has some vegetation.
3. Work should be undertaken to improve specification of the reference level where short vegetation is prescribed.
4. All model runs should be validated by measurements taken during combined model and field experiments.

5. Liaison should be established with the canopy program developers to insure updates and modifications to the algorithms are implemented as they occur.

REFERENCES

1. Atkinson, B.W. (1981): "Meso-scale atmospheric circulations," Academic Press, 495pp.

2. Bader, D.C. (1985): "Mesoscale Boundary Layer Development over Mountainous Terrain," Ph.D. Dissertation, Atmos. Sci. Paper No. 396, Colorado State University, Ft. Collins, Co. 80523, 1-251.

3. Banta, R.M (1986): Daytime boundary layer evolution over mountainous terrain. Part II: Numerical studies of upslope flow duration. Mo. Wea. Rev., 114: 1112-1130.

4. Banta, R., Pantley, K.C. and P.T. Gannon, Sr. (1987): Soil moisture and ridge height effects on katabatic flow simulations, Fourth Conference on Mountain Meteorology, Aug 25-28, 1987, Seattle, WA, American Meteorological Society, Boston, MA.

5. Camillo, P. and T. J. Schmugge (1981): A computer program for the simulation of heat and moisture flow in soils. AgRISTARS TM 82121, National Technical Information Service, 87pp.

6. Deardorff, J.W. (1978): Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation. J. Geophys. Res., 83: 1889-1902.

7. Fleagle, R.E. (1950): "A Theory of Air Drainage," Journal of Meteorology, 17, 227-232.

8. Gannon, P.T., Sr. (1978): Influence of earth surface and cloud properties on the south Florida sea breeze, NOAA Tech. Report ERL 402-NHEML 2, Environmental Research Laboratories, Boulder, Colorado 80303, 91 pp.

9. Gannon, P.T., Sr. (1986): Response of downslope and Florida mesoscale wind systems to physiographic features, Final Report, Summer Faculty Research Program/Graduate Student Summer Support Program, Air Force Geophysics Laboratory, Hanscom Air Force Base, MA, 14 pp.

10. Garrett, A.J. (1983): Drainage Flow Prediction with a One-Dimensional Model Including Canopy Soil and Radiation Parameterizations, Journal of Climate and Applied Meteorology, 22: 79-91.

11. Goudriann, J. (1977): Crop micrometeorology: a simulation Study. Simulation Monographs, Pudoc, Wageningen, 249pp.

12. Hiramatsu, Y., Takuro, S., and Maitani, T. (1984), Goudriaan's model of crop micrometeorology applied to

rice crop. "Berichte des Ohara Insituts fur landwirtschaftliche Biologie, Okayama Universitat" Bd. XVIII, Heft 4, S. 183-209, MARZ 1984.

13. Hootman, B.W. and W. Blumen (1983): "Analysis of Nighttime Drainage Winds in Boulder, Colorado during 1980," Monthly Weather Review, 111, 1052-1061.

14. Horst, T.W. and J.C. Doran (1986): "Nocturnal Drainage Flow on Simple Slopes," Boundary-Layer Meteorology, 34, 263-286.

15. Jialin, C. (1984): Mathematical Analysis and simulation of crop micrometeorology, agricultural University, Wageningen, The Netherlands, 116pp.

16. Louis, J. (1979): "A Parametric Model of Vertical Eddy Fluxes in the Atmosphere," Boundary Layer Meteorology, 17, 187-202.

17. Mahrt, L., and S. Larsen (1982): "Small Scale Drainage Front," Tellus, 34, 579-587.

18. Manins, P.C. and B.L. Sawford (1979): "Katabatic Winds: A Field Case Study," Quarterly Journal of the Royal Meteorological Society, 105, 1011-1025.

19. McCumber, M.C. (9180): A numerical simulation of the influence of heat and moisture fluxes upon mesoscale circulations. Ph.D. dissertation, University of Virginia, Charlottesville, VA 22903, 255pp.

20. McCumber, M.C. and R.A. Pielke (1981): Simulation of the effects of surface fluxes of heat and moisture in a mesoscale numerical model. Part I: Soil layer. J. Geophys. Res., 86: 9929-9938.

21. McNider, R.T. (1982): "A Note on Velocity Fluctuations in Drainage Flows," Journal of the Atmospheric Sciences, 39, 1658-1660.

22. McNider, R.T., and R.A. Pielke (1984): "Numerical Simulation of Slope and Mountain Flows," Journal of Climate and Applied Meteorology, 23, 1441-1453.

23. Petkovsek, F. and A. Hocevar (1971): "Night Drainage Winds," Archives for Meteorology, Geophysics and Bioclimatology, 20, 353-360.

24. Pielke, R.A. (1974): A three-dimensional numerical model of the sea breezes over south Florida. Mon. Wea. Rev., 102: 115-139.

25. Tremback, C. J., and R. Kessler (1985): A surface temperature and moisture parometerization for use in mesoscale numerical models. Seventh Conference on Numerical Weather Prediction, June 17-20, 1985, Montreal, Quebec, Canada, American Meteorological Society, Boston, MA.

26. Tripoli, G.J. and W.R. Cotton (1982): "The Colorado State University Three-Dimensional Cloud/Mesoscale Model-1982. I. General Theoretical Framework and Sensitivity Experiments," Journal de Recherches Atmospheriques, 16, 185-219.

27. Yamada, T. (1985): "Numerical Simulations of the Night 2 Data of the 1980 ASCOT Experiments in the California Geysers Area," Archives for Meteorology, Geophysics and Bioclimatology, 34, 223-247.

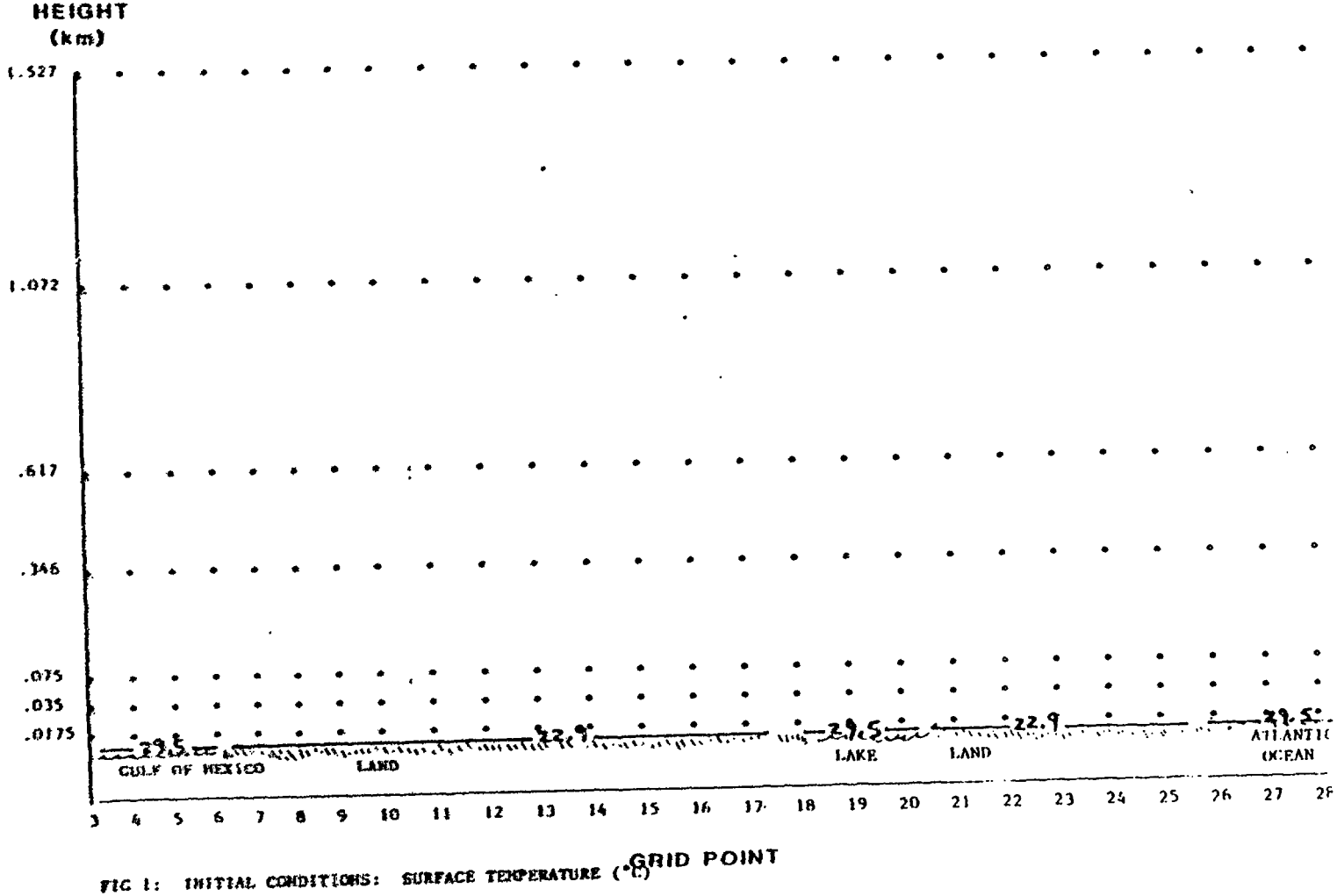


FIG 1: INITIAL CONDITIONS: SURFACE TEMPERATURE (°C)

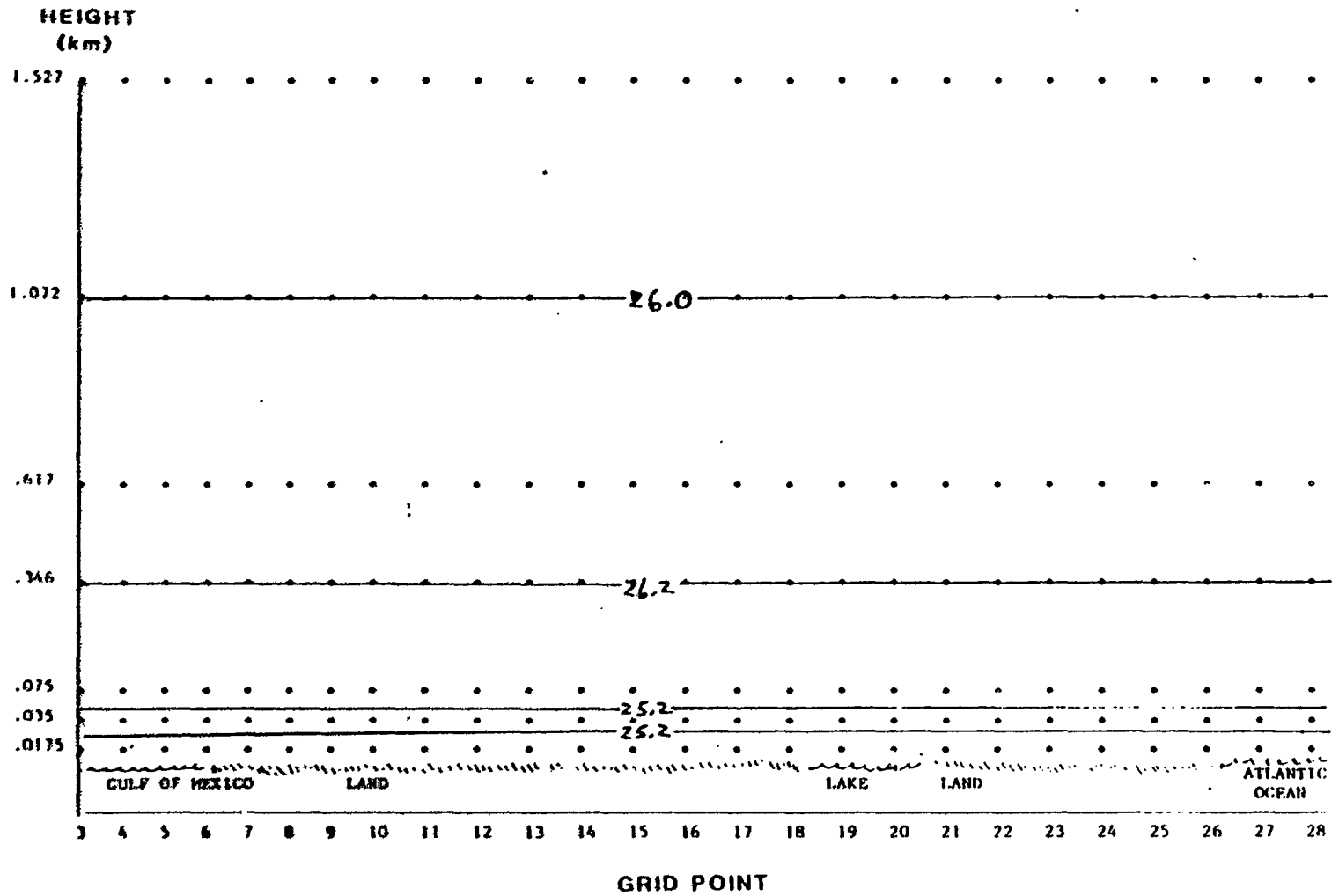


FIG 2: INITIAL CONDITIONS: POTENTIAL TEMPERATURE (°C)

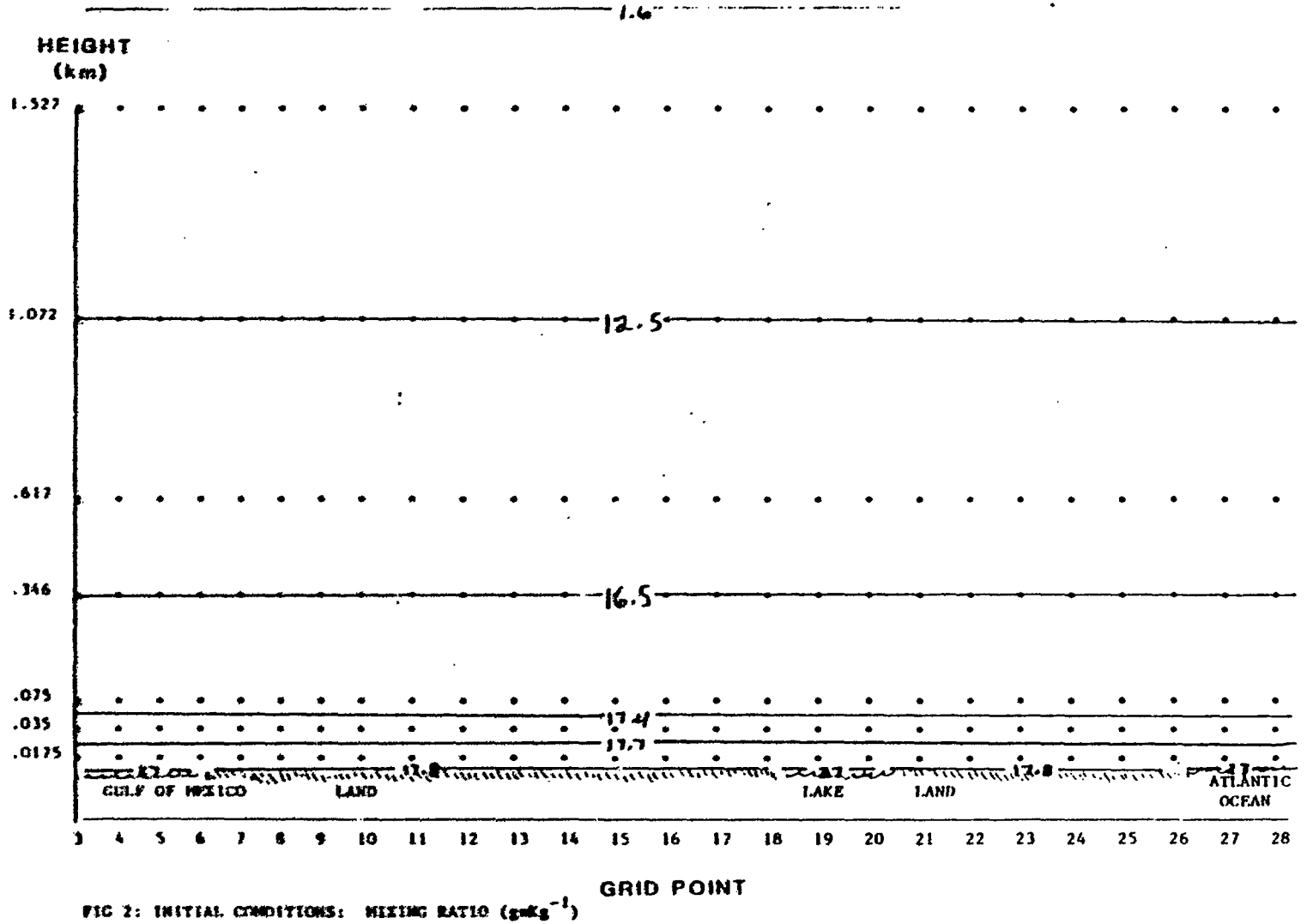


FIG 2: INITIAL CONDITIONS: MIXING RATIO ($gmKg^{-1}$)

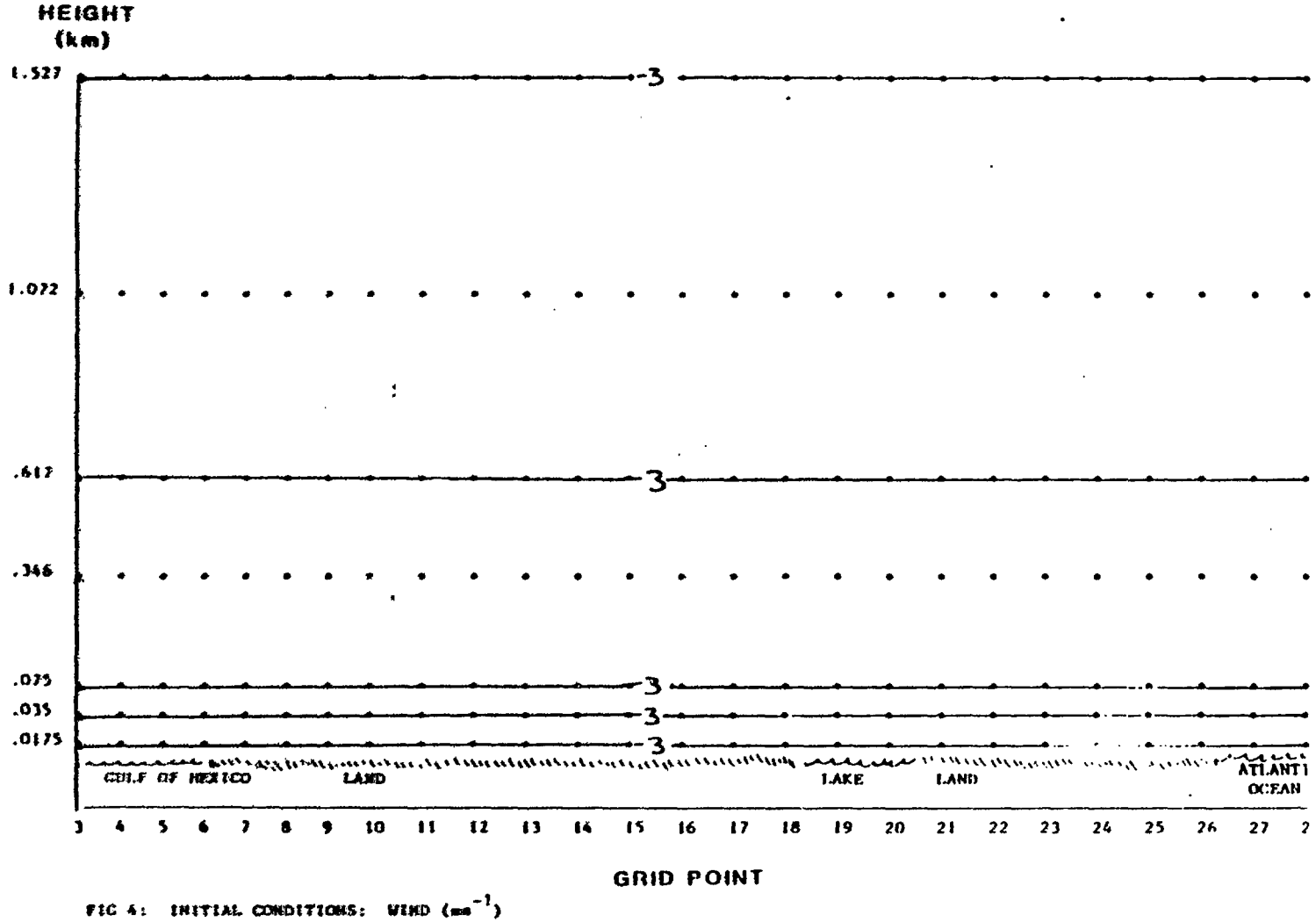


FIG 4: INITIAL CONDITIONS: WIND (ms^{-1})

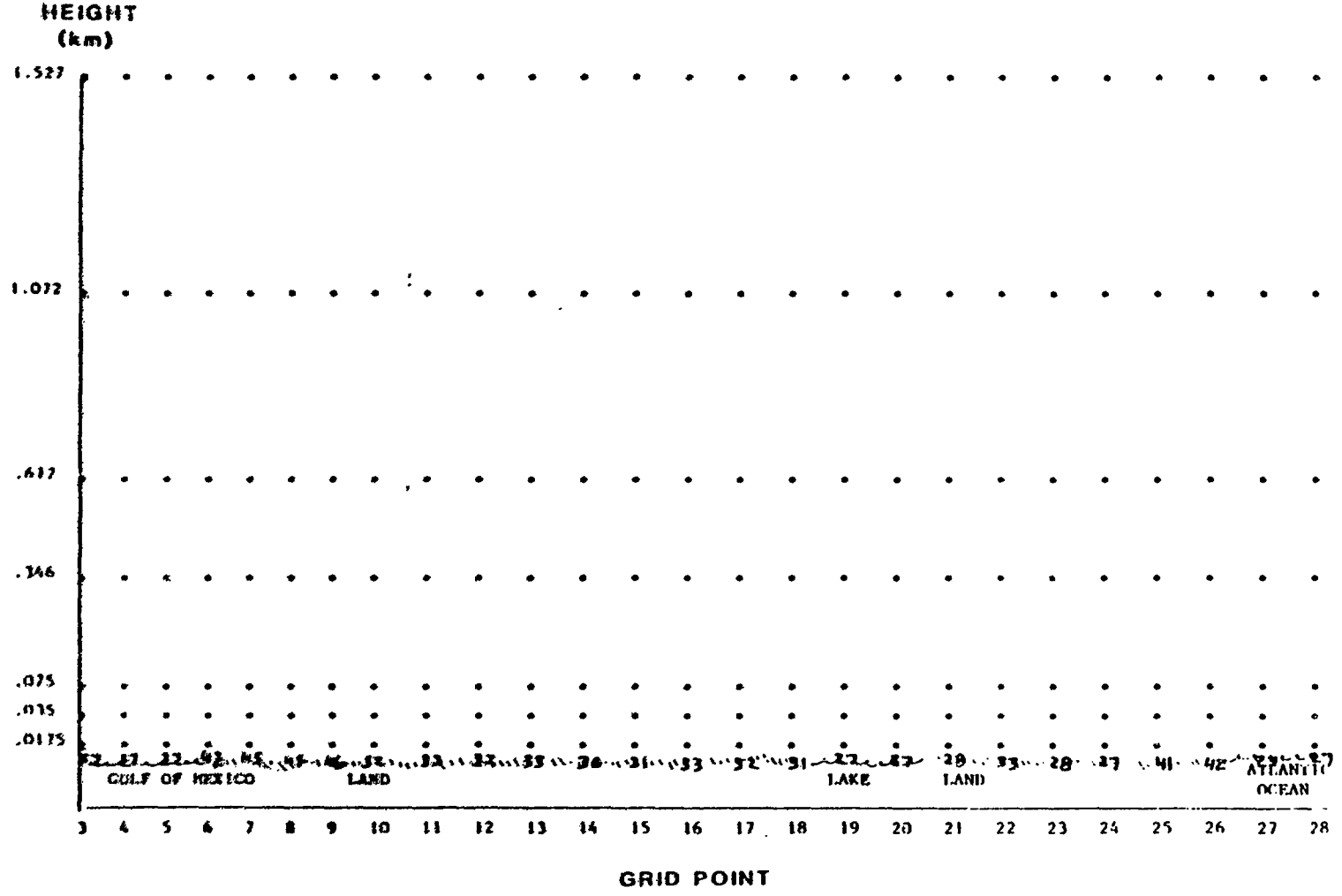


FIG 5: SURFACE TEMPERATURE (°C) AT 1600 LST, EXP 1

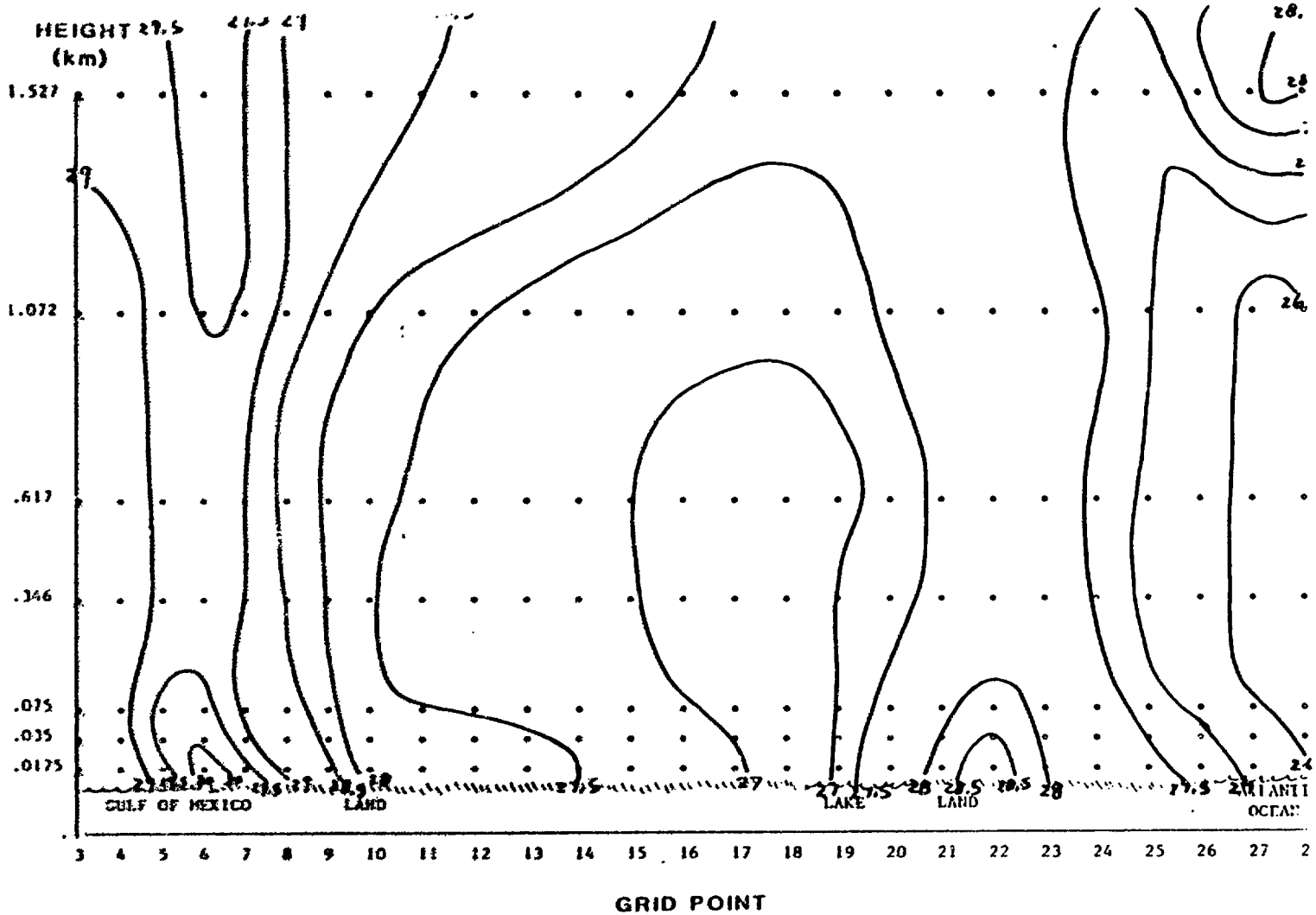


FIG 6: POTENTIAL TEMPERATURE (°C) AT 1600 LST, EXP 1

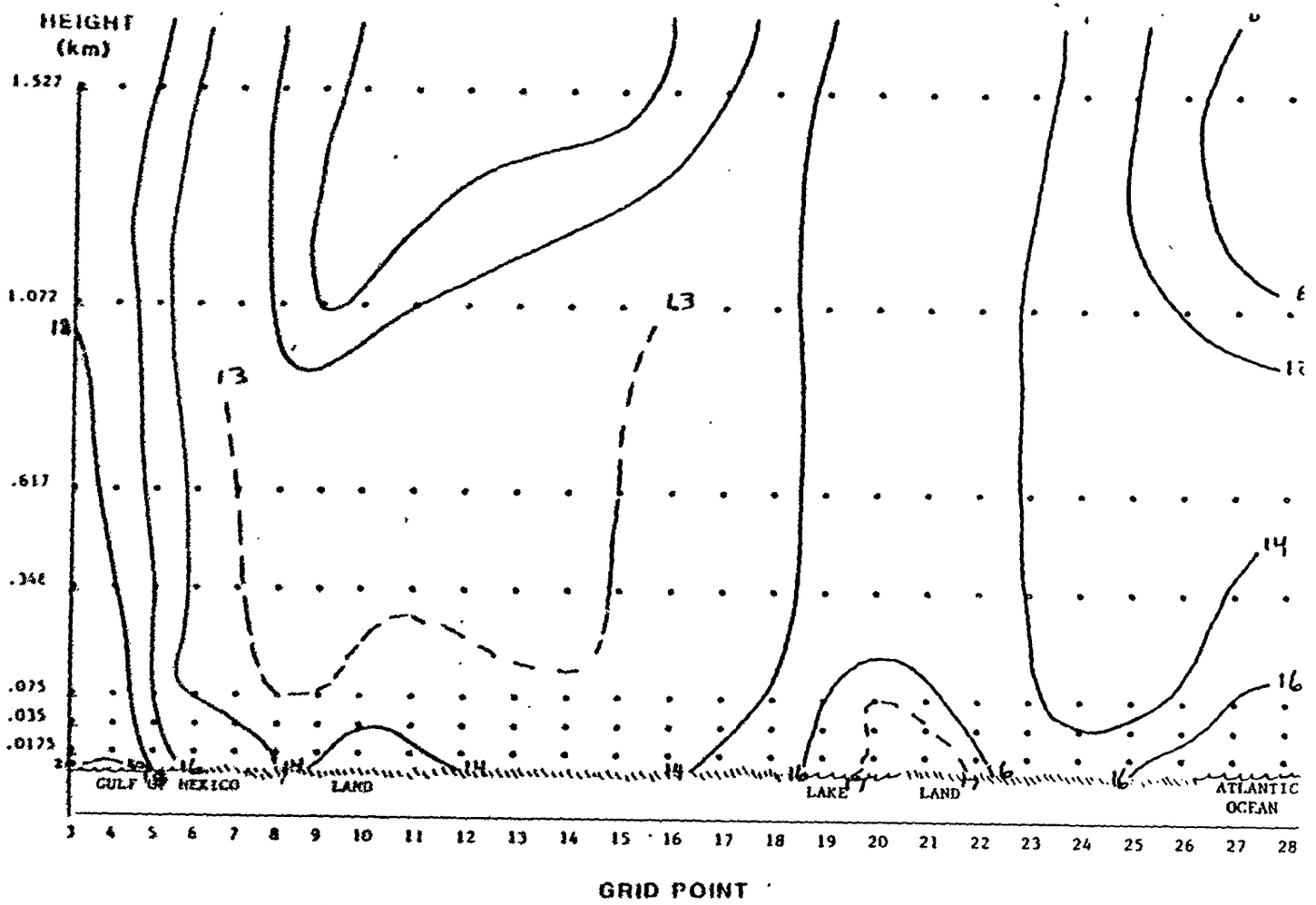


FIG 7: MOISTURE ($gmKg^{-1}$) AT 1600 LST, EXP 1

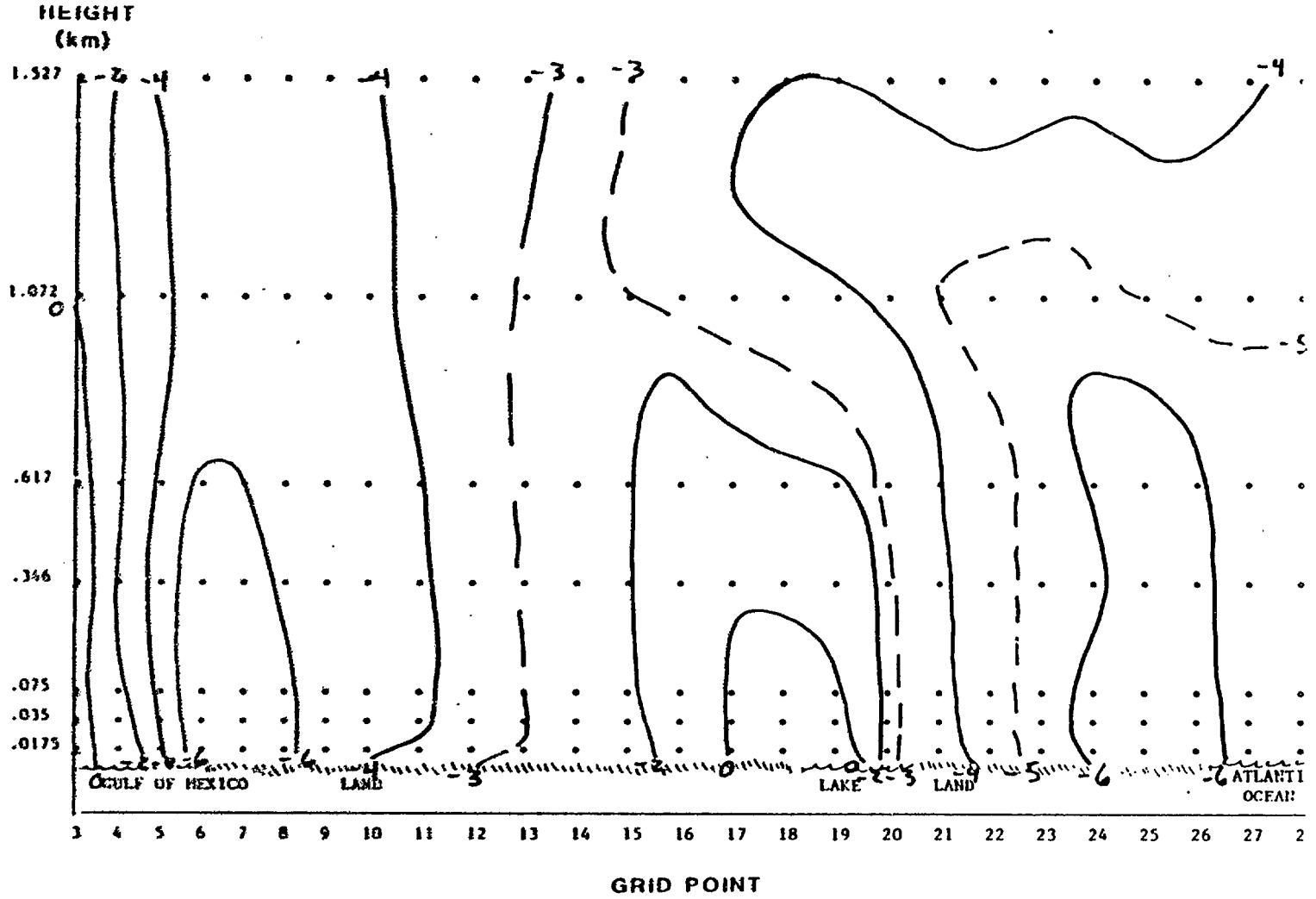


FIG 8: WIND (ms^{-1}) AT 1600 LST, EXP 1

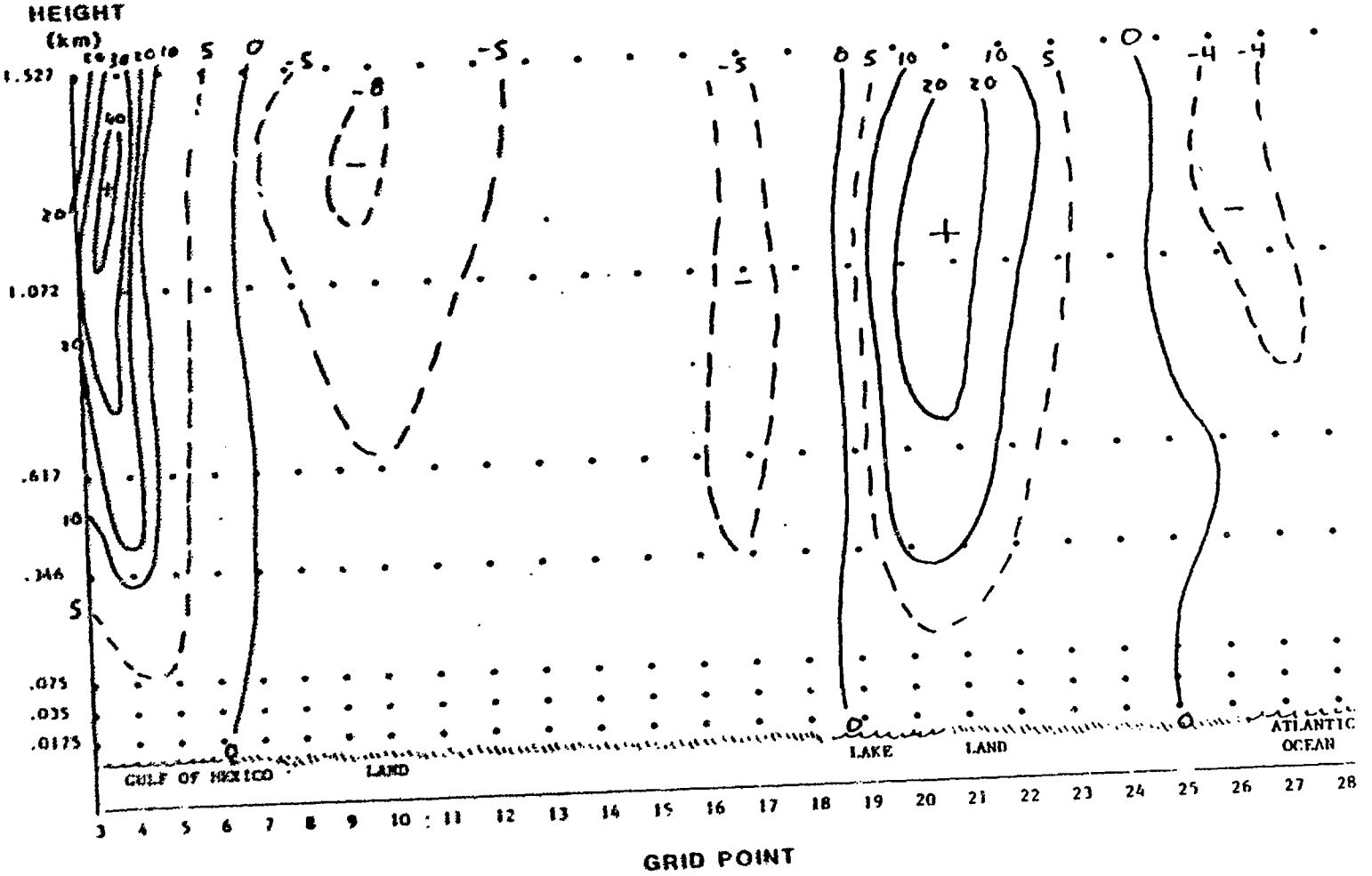


FIG 9: VERTICAL MOTION (cms^{-1}) AT 1600 LST, EXP 1

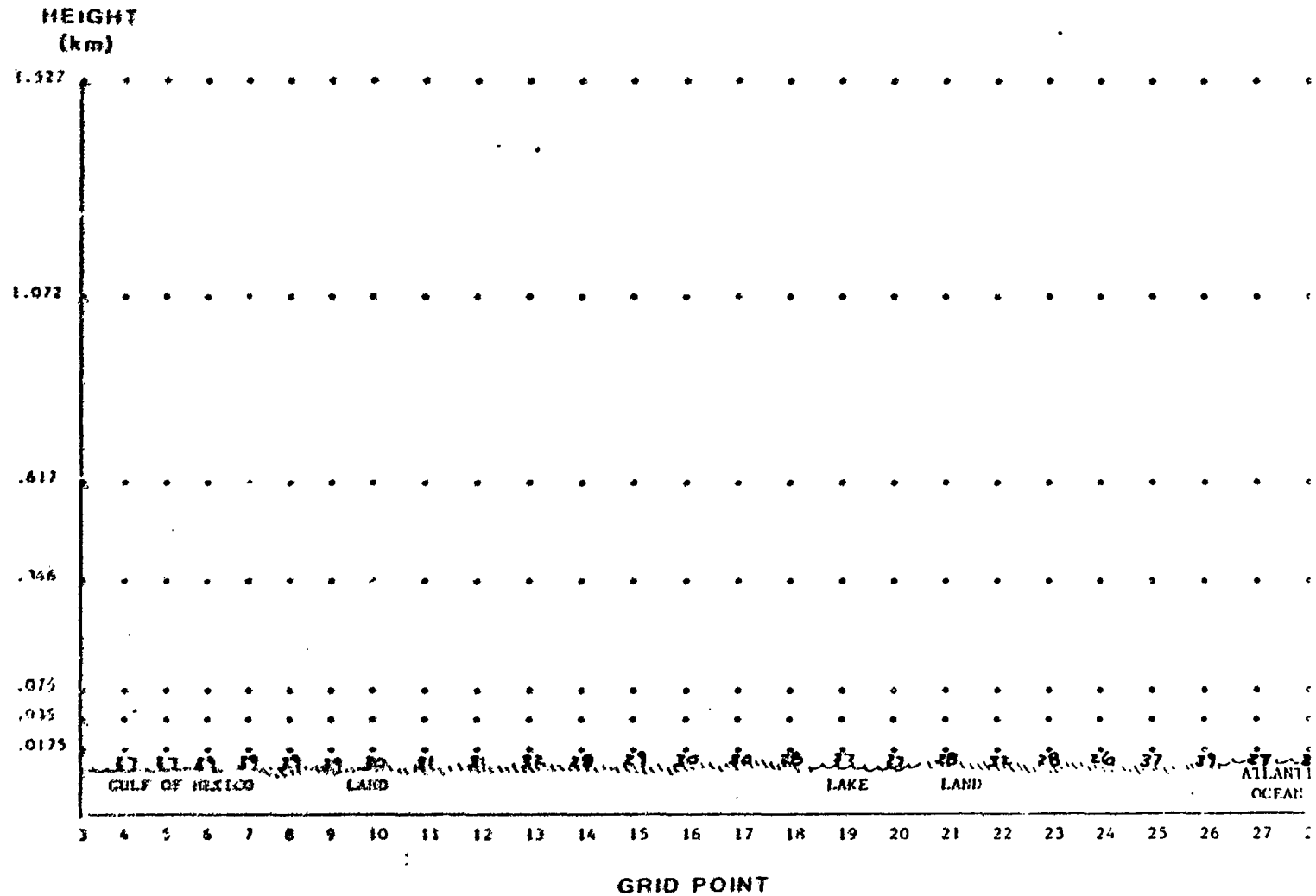


FIG 10: SURFACE TEMPERATURE (°C) AT 1600 LST, EXP 2

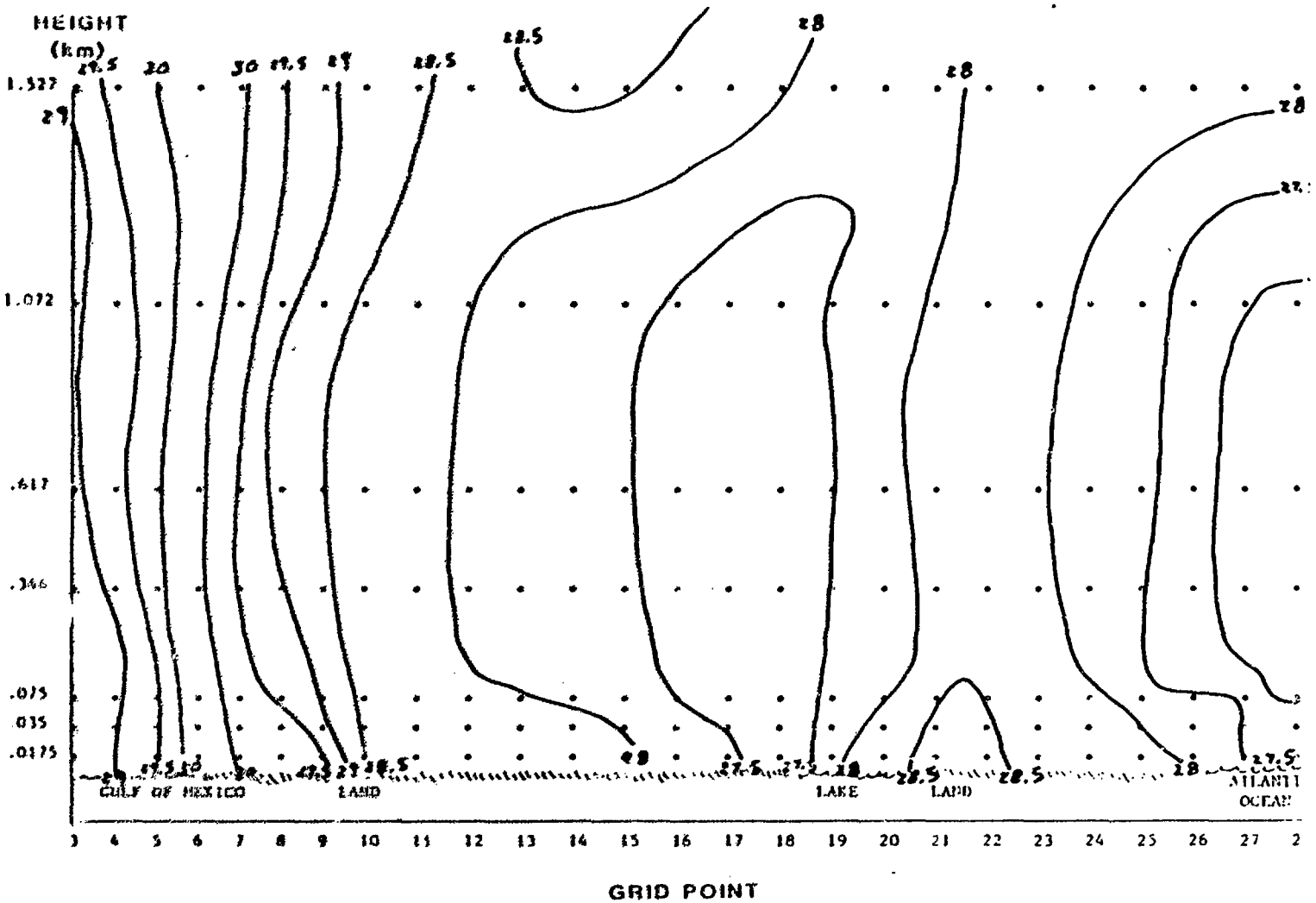


FIG 11: POTENTIAL TEMPERATURE (°C) AT 1600 LST, EXP 2

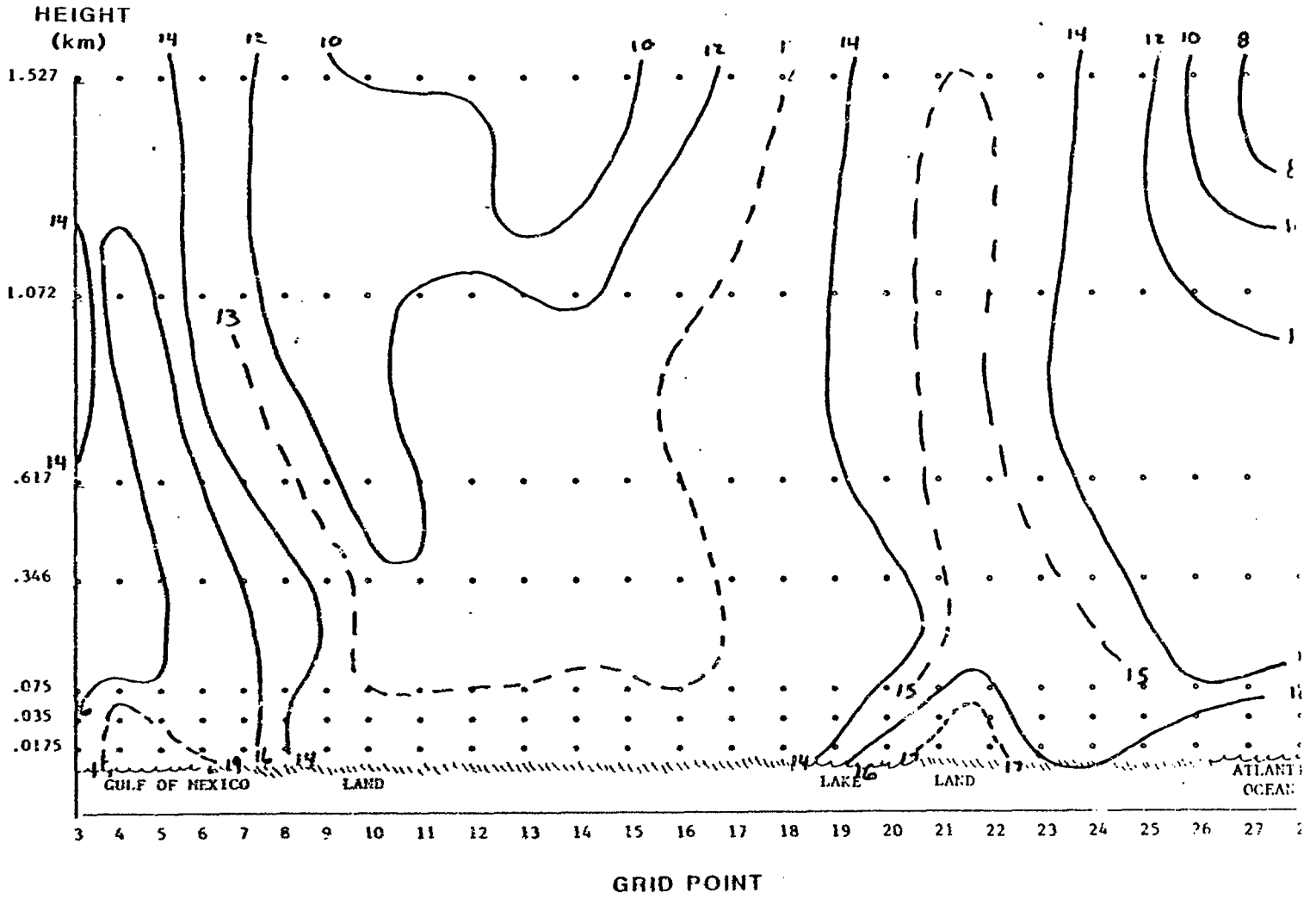


FIG 12: MOISTURE ($gmKg^{-1}$) AT 1600 LST, EXP 2

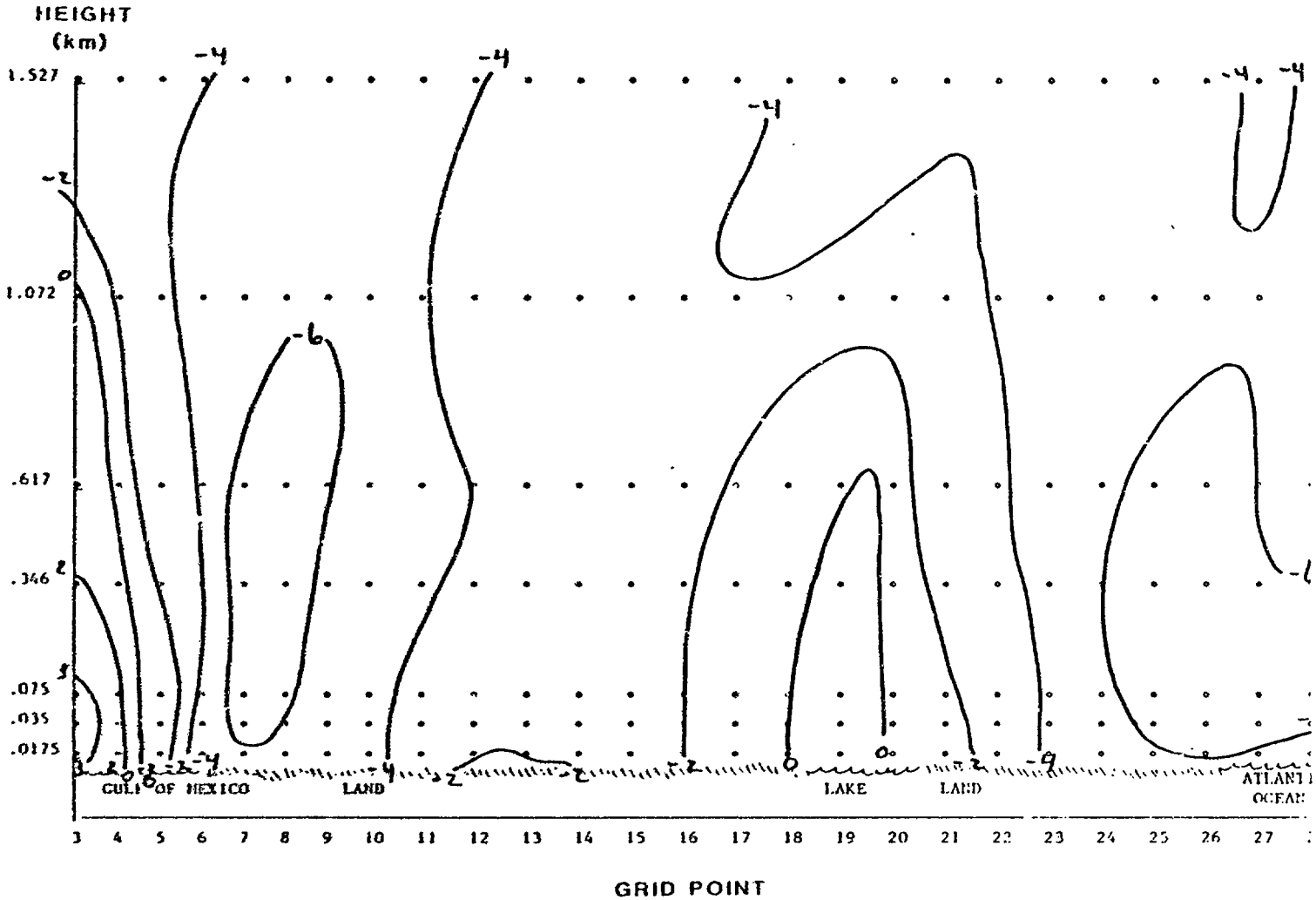


FIG 13: WIND (ms^{-1}) AT 1600 LST, EXP 2

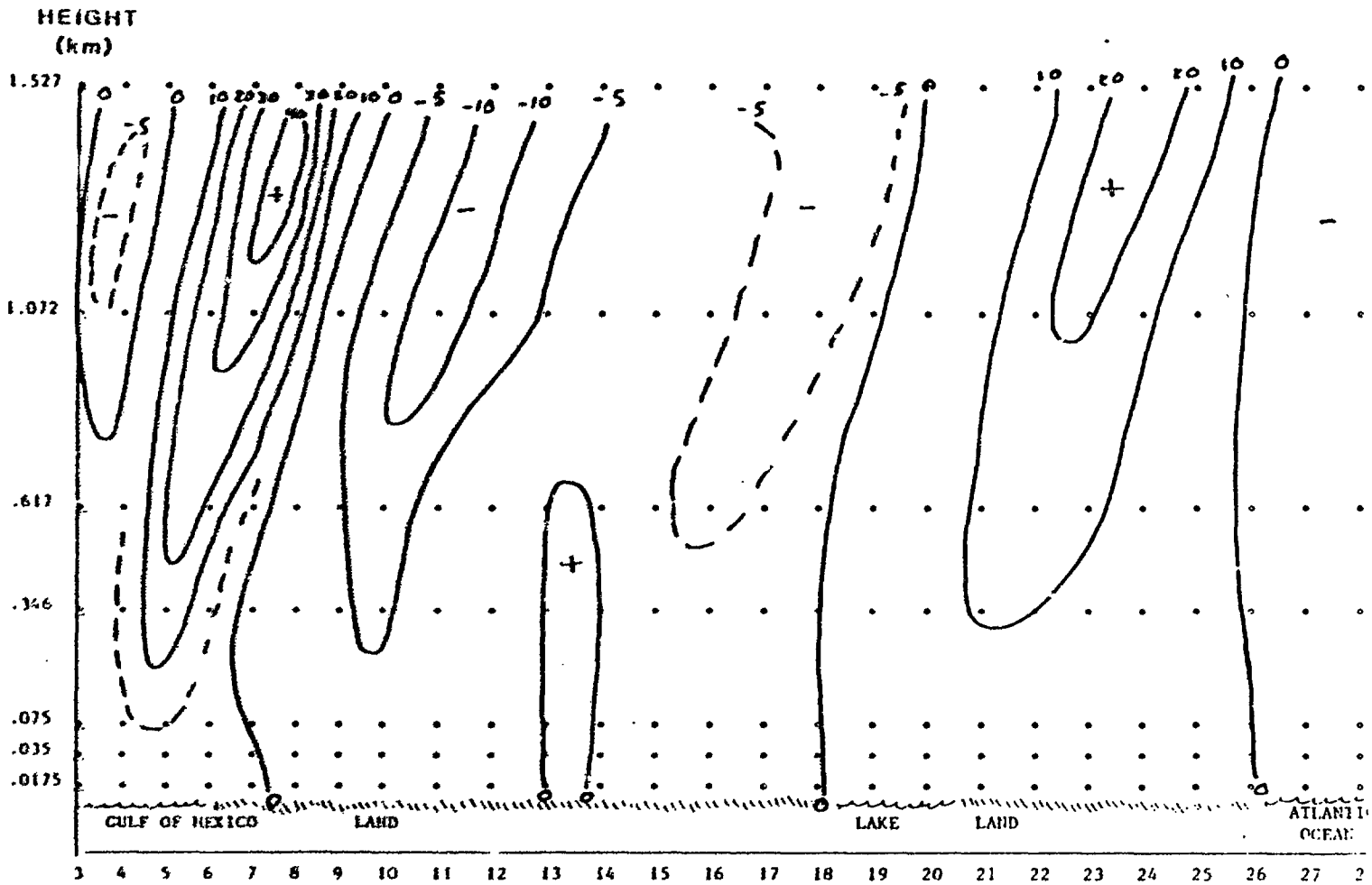


FIG 14: VERTICAL MOTION (cms^{-1}) AT 1600 LST, EXP 2

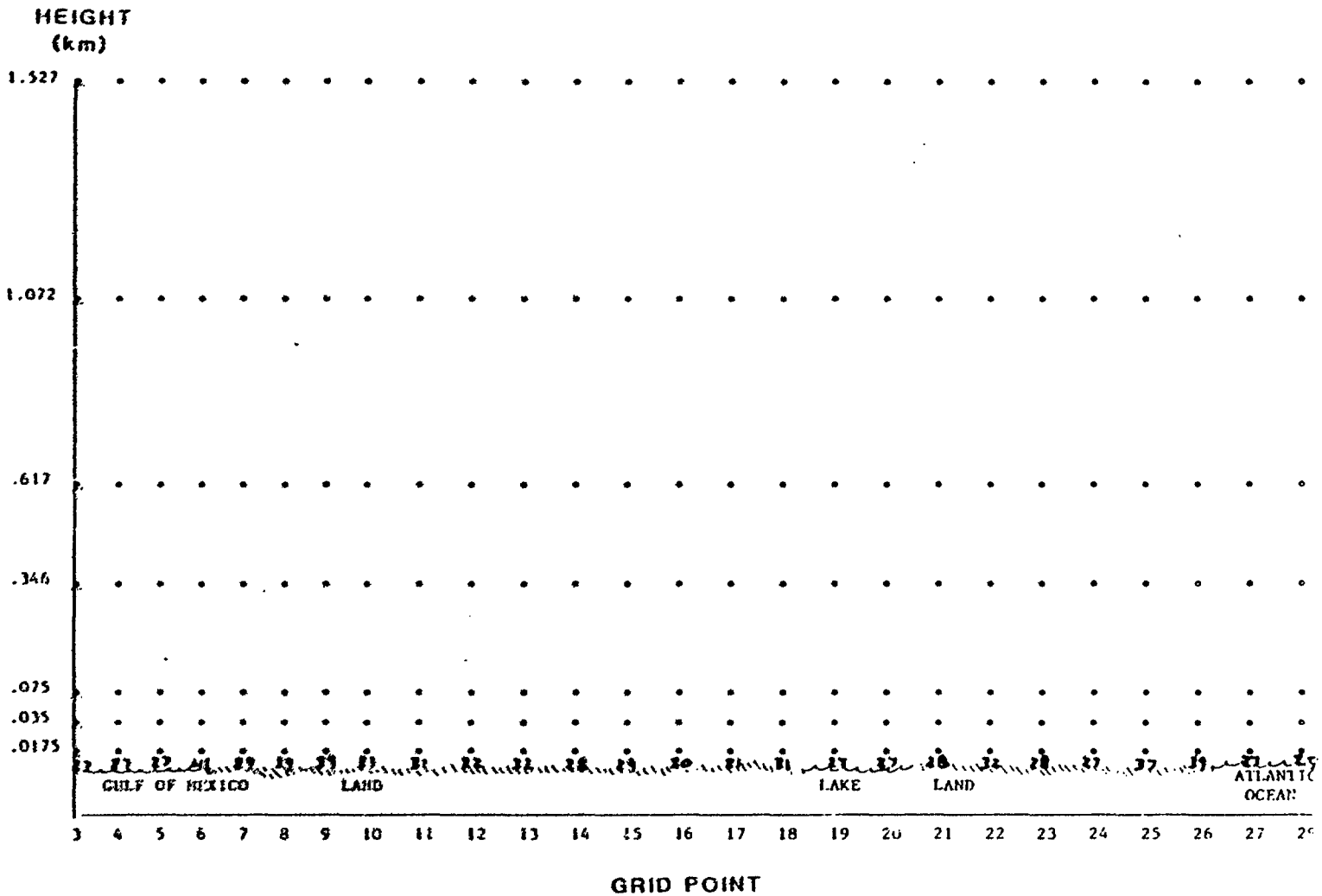


FIG 15: SURFACE TEMPERATURE (°C) AT 1600 LST, EXP 3

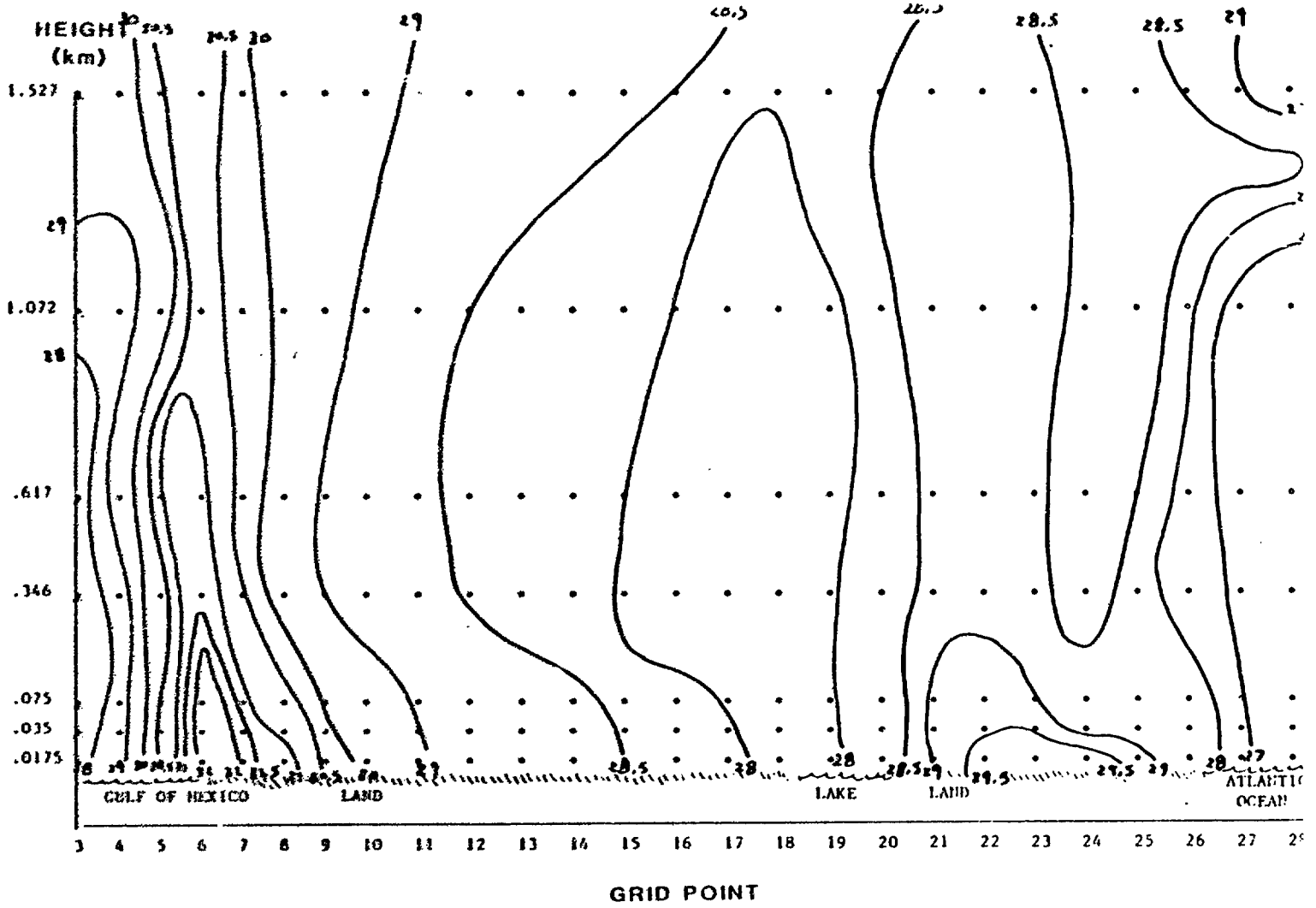


FIG 16: POTENTIAL TEMPERATURE (°C) AT 1600 LST, EXP 3

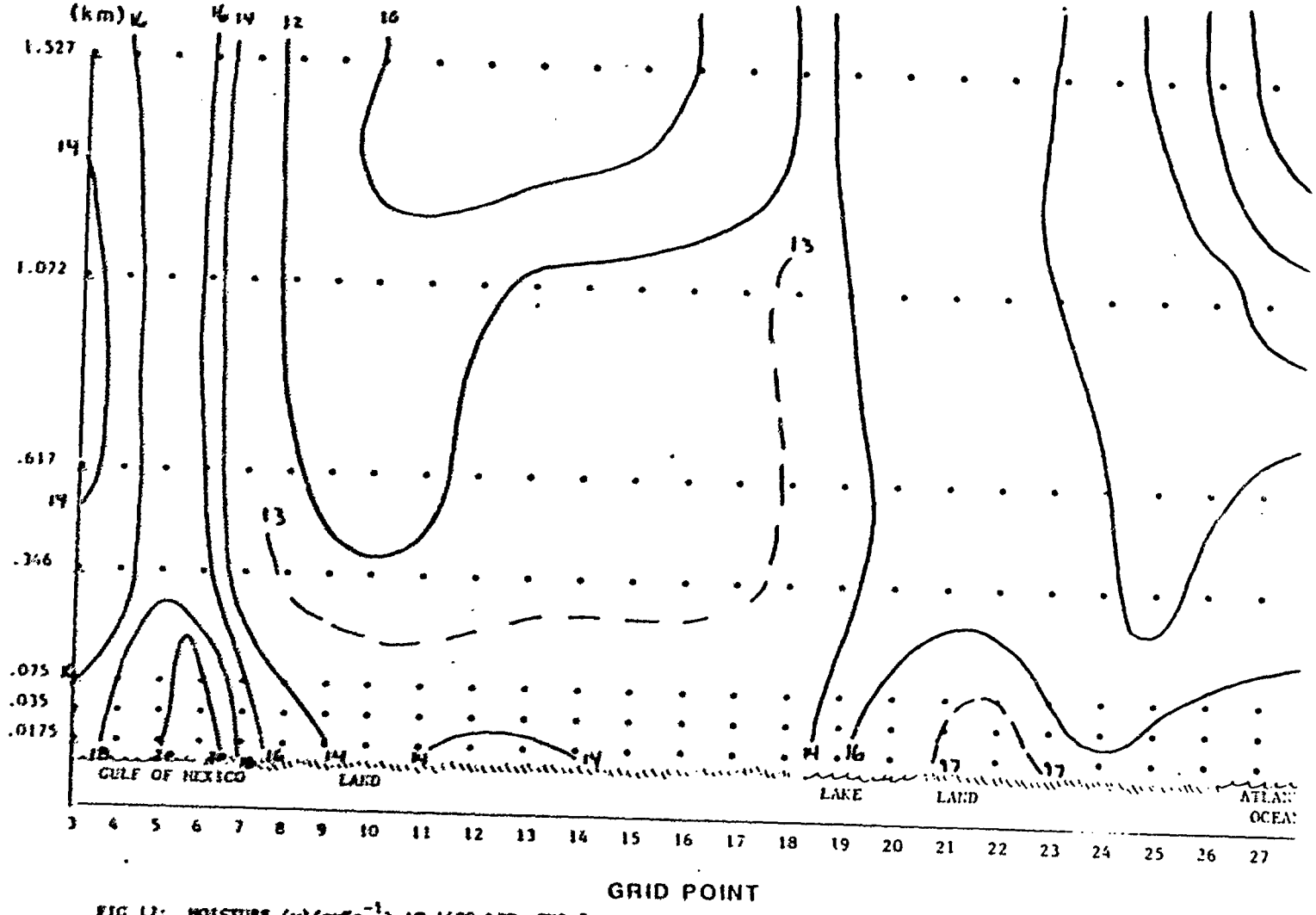


FIG 17: MOISTURE (w)(g^{-1}) AT 1600 LST, EXP 3

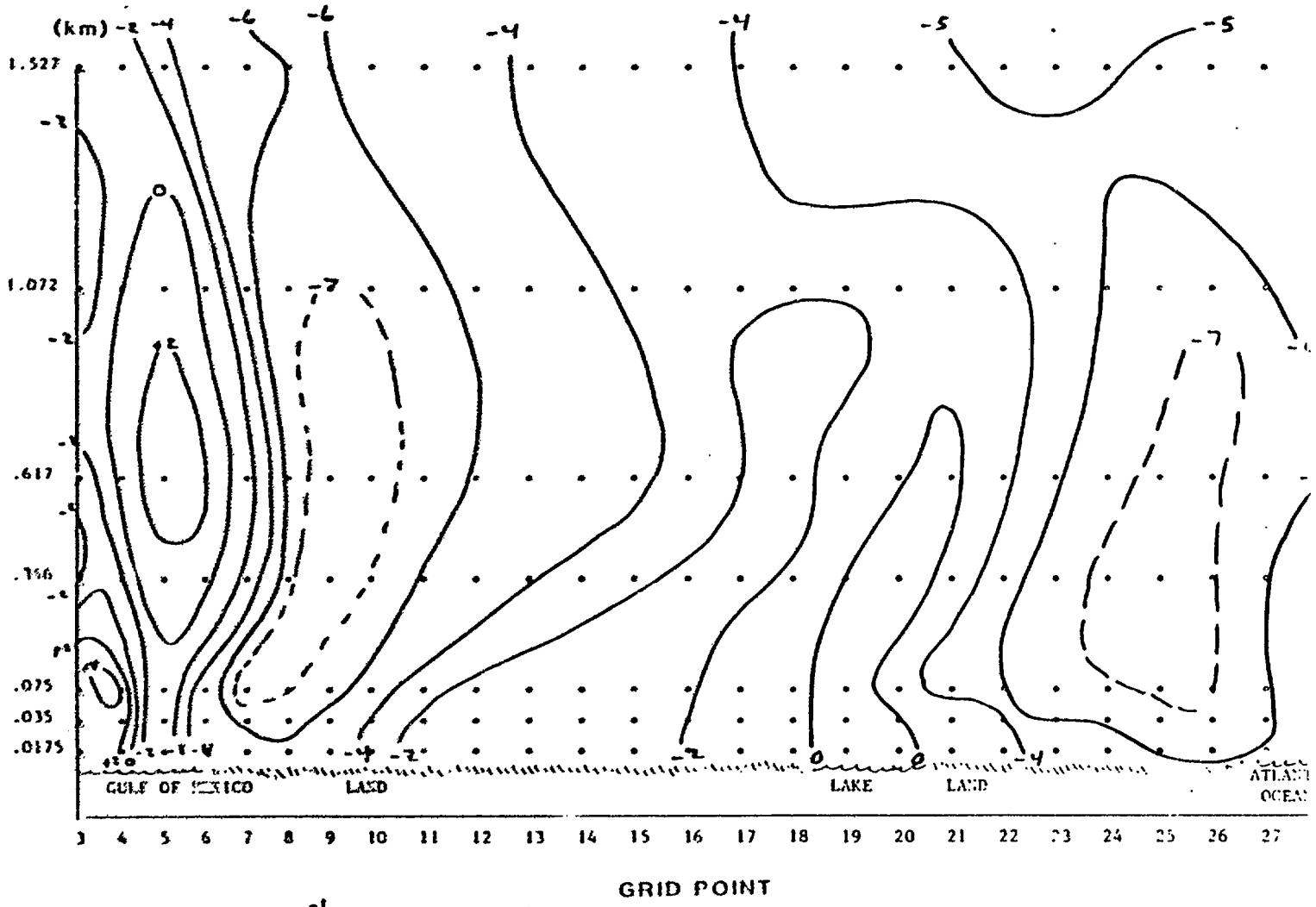


FIG 18: WIND (u)(ms^{-1}) AT 1600 LST, EXP 3

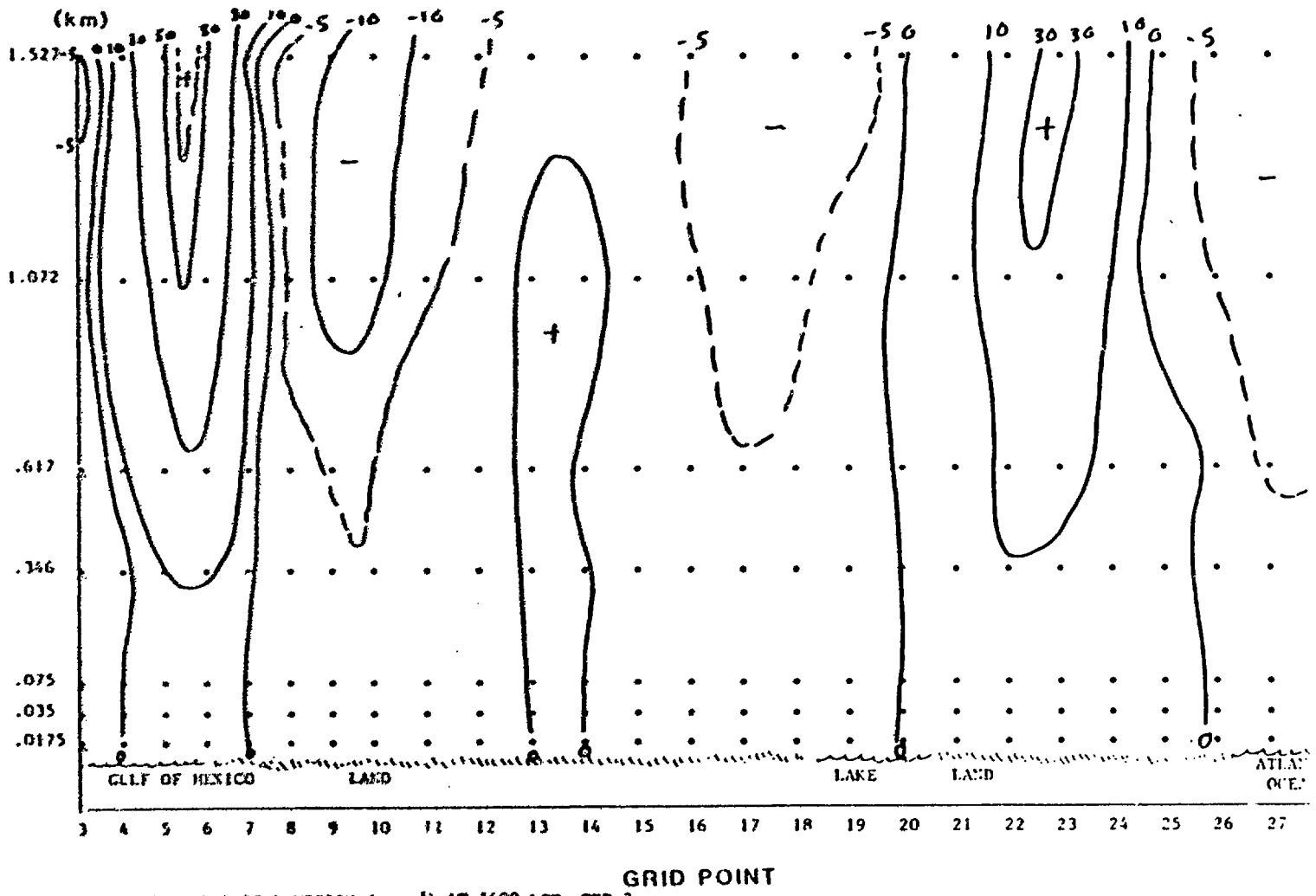


FIG 19: VERTICAL MOTION (cms^{-1}) AT 1600 LST, EXP 3

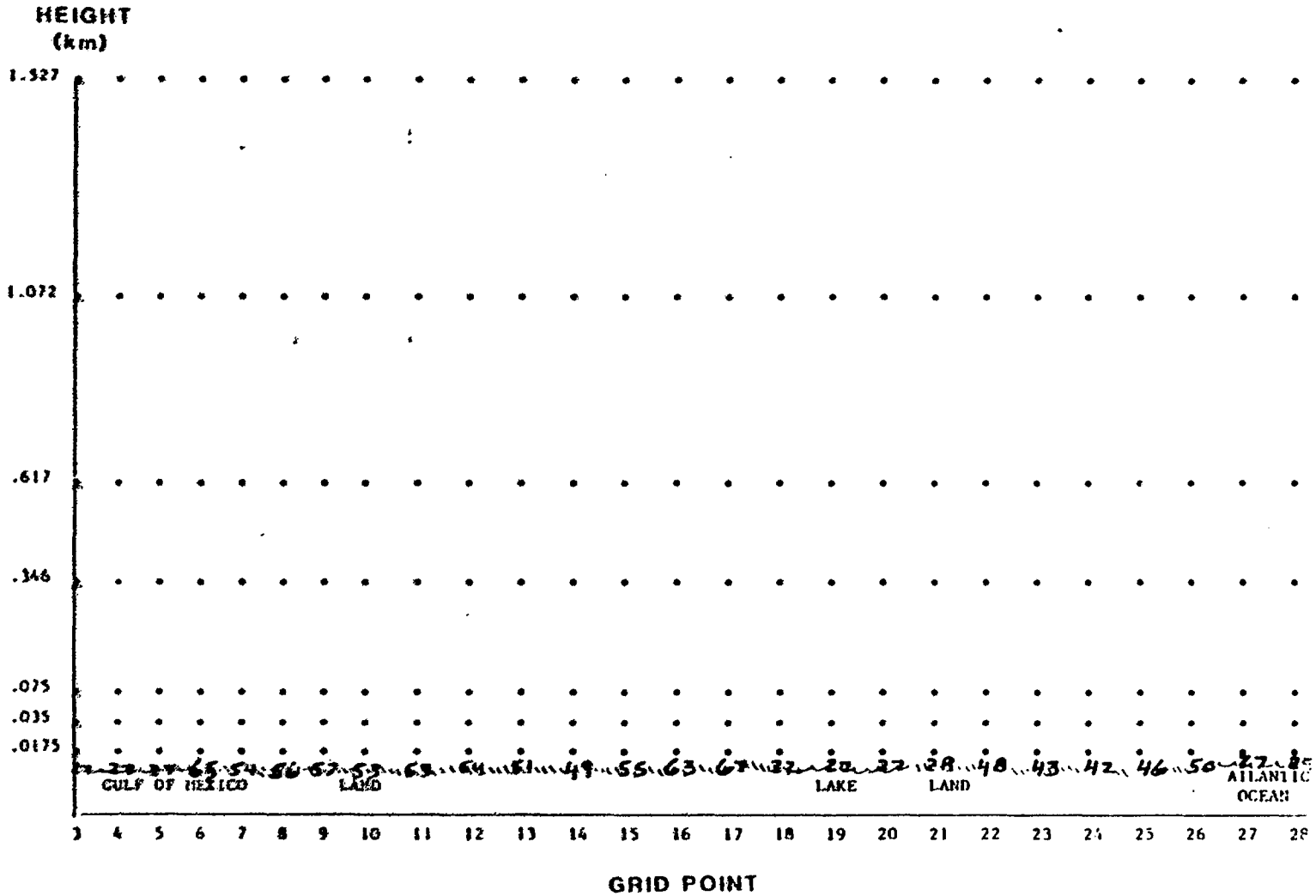


FIG 20: SURFACE TEMPERATURE (°C) AT 1600 LST, EXP 4

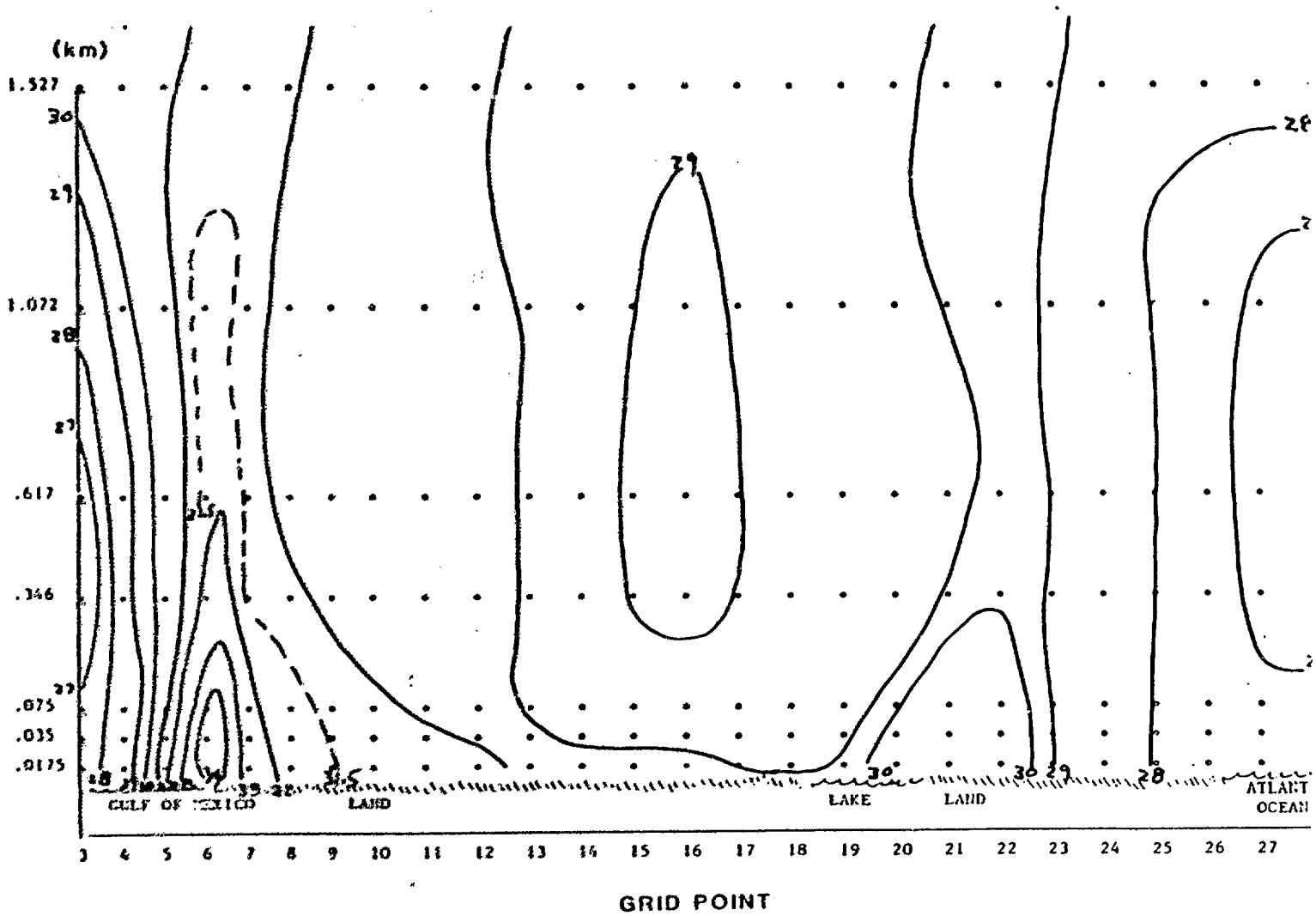


FIG 21: POTENTIAL TEMPERATURE (°C) AT 1600 LST, EXP 4

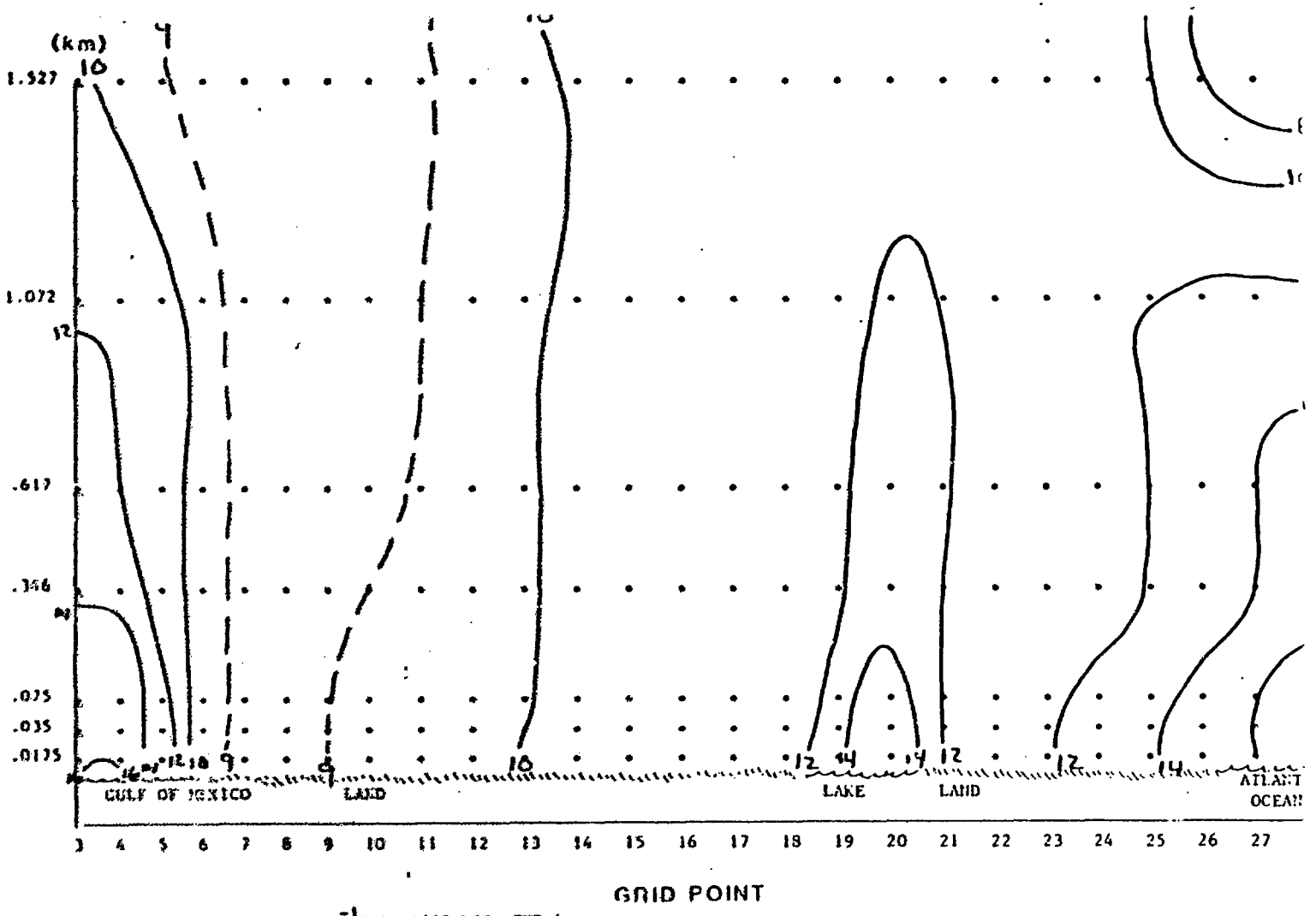


FIG 22: MOISTURE (gm/g^{-1}) AT 1600 LST, EXP 4

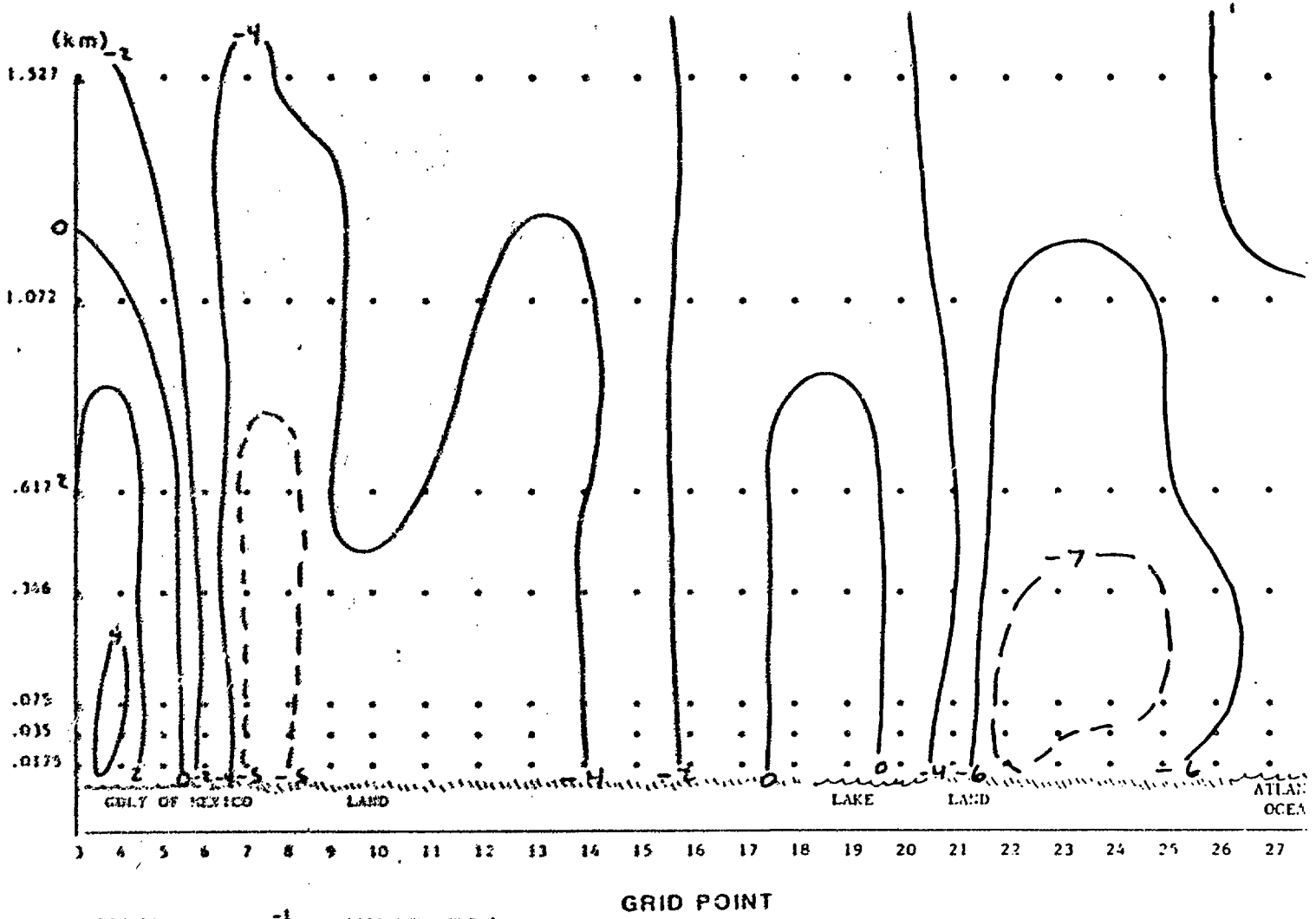


FIG 23: WIND (ms^{-1}) AT 1600 LST, EXP 4

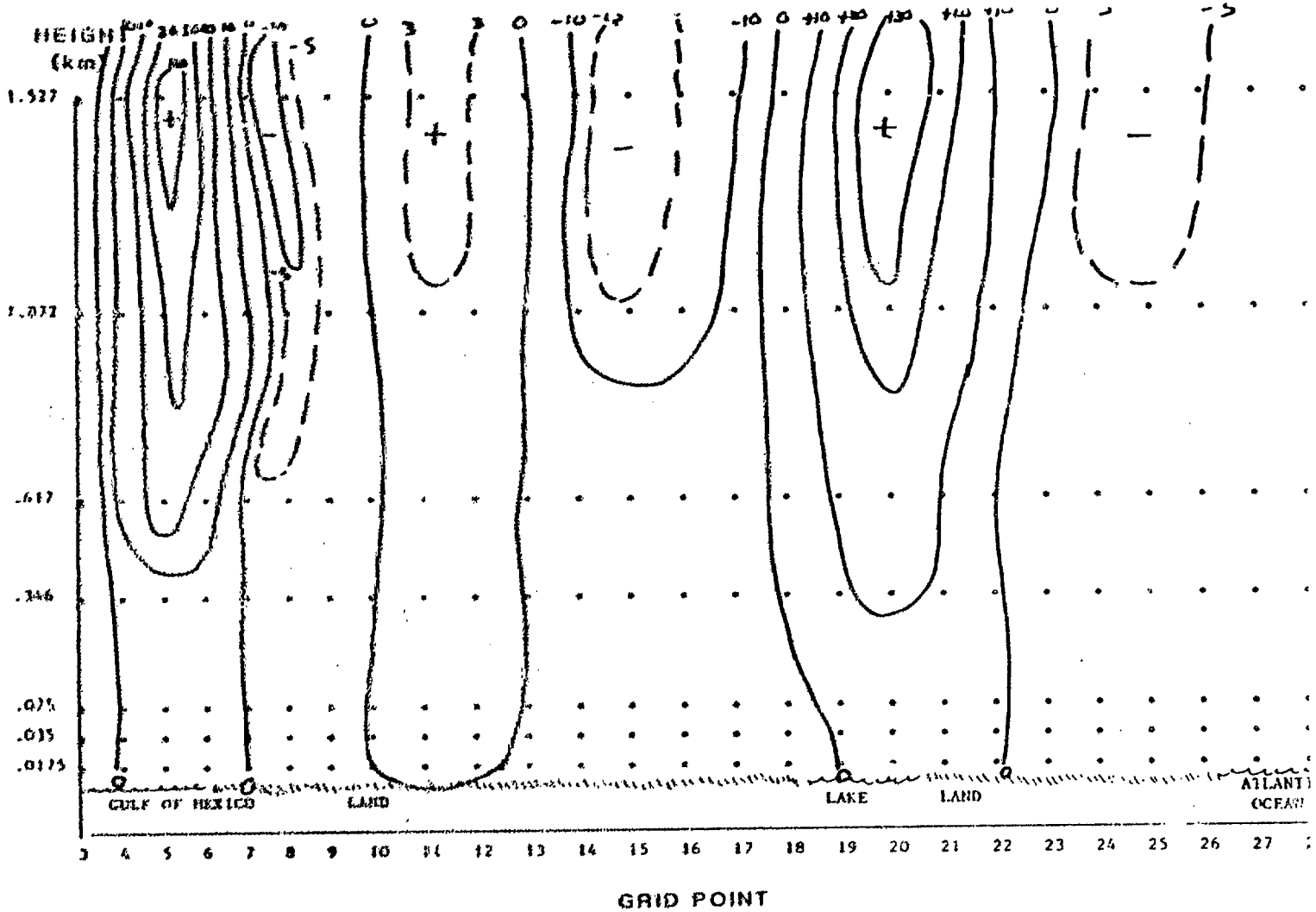


FIG 24: VERTICAL MOTION (cms^{-1}) AT 1600 LST, EXP 4

ACKNOWLEDGMENTS

I am grateful for this second research opportunity provided me by the Air Force Office of Scientific Research. I am again grateful to Dr. Robert Banta for professional interactions and an opportunity to analyze and diagnose numerical model simulations of katabatic flows. Special thanks go to S Sgt Charles Crouch of the Air Force Geophysics Laboratory for many hours spent converting computer tapes. Without the computer and programming expertise of Ms. Abigail Thomas, Coordinator of Academic Computing at Lyndon State College, this work would not have been successfully completed. I am indebted to Dr. I. R. Goudriaan of Landbouwuniversiteit, Wageningen, The Netherlands for providing guidance and the canopy algorithms of Dr. Chen Jialin. Dr. Craig Tremback and Capt Kim Pantley assisted at critical points during this project. I appreciate the administrative and facilities support provided me by President Clive Veri and Academic Dean Perry Viles of Lyndon State College. I am grateful to Ms. Susan Englese for her work in preparing and coordinating the proposal to Universal Energy Systems, Inc. Ms. Sheryl Laing provided excellent administrative and typing support.

Finally, I would like to thank my wife, Donna, for her support and patience during the past year.

Appendices can be obtained from

Universal Energy Systems, Inc.

RESEARCH INITIATION PROGRAM

Sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the

Universal Energy Systems, Inc.

FINAL REPORT

EFFECTS OF EXERCISE AND DOBUTAMINE ON SUSPENSION HYPOKINESIA/HYPODYNAMIA DECONDITIONING IN RATS

Prepared by: Beverly Elaine Girten

Departments: Exercise Physiology
Physiological Chemistry
Veterinary Physiology

University: The Ohio State University

Grant Period: December, 1986 - December, 1987

EFFECTS OF EXERCISE AND DOBUTAMINE ON SUSPENSION HYPOKINESIA/HYPODYNAMIA DECONDITIONING IN RATS

by

Beverly Elaine Girten

ABSTRACT

Two potential interventions that may have positive effects on deconditioning changes in general, are exercise conditioning (24, 29, 43) and administration of dobutamine, a drug that has been shown to induce or maintain conditioning effects similar to those seen in exercise training (42, 44, 72). The present study utilized a rat hindlimb suspension model in an attempt to determine if an exercise training program or the administration of dobutamine, or a combination of both, could have a positive effect on deconditioning changes. Basic parameters such as weight, food consumption, activity levels, organ mass, muscle mass, heart rate and electrocardiographic waveform response of the four sedentary groups were measured during the summer research project. Additional factors that provide insight into biochemical adaptations and alterations in metabolic control factors and cardiovascular response and control were included among items analyzed for all groups in the current proposal.

Many of the results obtained during this investigation suggest that exercise training prior to suspension hypokinesia/hypodynamia can be helpful. Additionally, even though the 14 day dobutamine treatment did not induce tremendous changes in and of itself, it did appear beneficial in maintaining some of the positive effects induced by exercise training prior to the suspension. The most notable exception to this was seen in the plasma somatomedin-C.

ACKNOWLEDGMENTS

I would like to thank the Air Force Systems Command, the Air Force Office of Scientific Research and the Armstrong Aerospace Medical Research Laboratory for sponsorship of this research. Sincere appreciation and thanks are extended to Leon Kazarian, Robert Hamlin, A. John Merola, David Lamb, Mike Sherman, Barb Smith, and, in particular Clarence Oloff, for their respective contributions to this project. In addition, I would also like to acknowledge and thank Emily Morey-Holton (NASA-Ames) for providing advice and details concerning the hindlimb suspension model and for her willingness to allow us to borrow the x-y axis suspension apparatus that were so critical to the study. The efforts of Pat Roberts and the UDRI group with regard to the construction of the suspension cages were also very much appreciated. Special thanks are also extended to all those who provided help and support along the way. Included among those from Wright-Patterson are Marvin Souder, Ed Eveland, Suzanne Smith, James Cooper, Pat Lewandowski, Nadia Greenidge, Gerald Harris and Brian Jenkins. Susan Bloomfield and Peggy Plato did a great deal to coordinate efforts at Ohio State and included among the others who helped at Ohio State were Marge Maxwell, Jamie Hobson, Steve Boggs, Andy Doyle, Art Dernbach, Dottie Sutter, Jim Schlub, Ray Boddour, Mark Rechsteiner, Lisa Carl, Ni Rong and Jiaping Gao. I would also like to thank my mother, Vedna Girten, and the rest of my family, for their constant support and encouragement during all of my research endeavors.

Introduction and Objectives

Alterations in the musculoskeletal and cardiovascular systems as a result of the near-weightless environment of space or in simulated weightless conditions are well documented (6, 7, 24, 43, 63). Although general information about the basic physiological and pathological effects of the deconditioning that results from the relative hypodynamia and hypokinesia is abundant (7, 24), there is still much to be learned. Investigation of effective countermeasures that might prevent or minimize the alterations that result from deconditioning are being restricted by lack of information on the basic biological mechanisms involved in these changes. Attempts to integrate observed changes with additional information on the mechanisms and control factors responsible for the change will help clarify and define possible solutions to problems that currently remain unresolved.

Two potential interventions that may have positive effects on deconditioning changes in general, are exercise conditioning (24, 29, 43) and administration of dobutamine, a drug that has been shown to induce or maintain conditioning effects similar to those seen in exercise training (42, 44, 72). The present study utilized a rat hindlimb suspension model in an attempt to determine if an exercise training program or the administration of dobutamine, or a combination of both, could have a positive effect on deconditioning changes. Basic parameters such as weight, food consumption, activity levels, organ mass, muscle mass, heart rate and electrocardiographic waveform response of the four sedentary groups were measured during the summer research project. Additional factors that provide insight into biochemical adaptations and alterations in metabolic control factors and cardiovascular response and control were included among items analyzed for all groups in the current proposal.

The primary objective of this study was to investigate the effect of exercise training and dobutamine on suspension hypokinesia/

hypodynamia (H/H) deconditioning by examining factors which reflect basic mechanisms of action related to the treatments. These factors provide information that account, in part, for observed changes (see Appendix B "Preliminary Data") in the cardiovascular and musculoskeletal systems of the experimental sedentary rats utilized during the initial summer research effort (see attached "Study Protocol," Appendix A) at Wright-Patterson Air Force Base. Investigation into the factors underlying the observed results required detailed analyses of a variety of tissues from the original eight groups of experimental animals, as well as utilization of a small number (24) of additional rats. Specific areas investigated included:

(1) biochemical adaptations that focused on oxidative and antioxidant enzyme change in skeletal muscle, but also examined myosin heavy chain isoenzyme patterns in cardiac muscle.

(2) metabolic control factors measuring alterations in levels of somatomedin C in blood and muscle, as well as glycogen, hexokinase and glycogen synthase in skeletal muscle.

(3) cardiovascular response and control that involved quantification of electrocardiographic (ECG) waveform changes, left ventricular wall area and examination for fibrosis. Initial investigations into measuring baroreceptor sensitivity and its role in control of heart rate and blood pressure response to the experimental treatments were also completed.

In addition to the above areas of study, alterations in the skeletal system of the experimental animals is also being investigated. These analyses will be completed at the Armstrong Aerospace Medical Research Laboratory (AAMRL) at Wright-Patterson Air Force Base, under the direction of Dr. Leon Kazarian and Mr. Clarence Oloff. Analyses

will include detailed histomorphometric investigations as well as mechanical testing of select bone. Work on the histomorphometric investigations have been started by S. Bloomfield with most of the work being done in Dr. Weisbrode's laboratory at The Ohio State University. Integration and synthesis of bone data with that which has been completed for this report will provide valuable information on the interactive nature of muscle and bone during conditions that simulate the weightless environment of space. Moreover, these data, when examined with regard to the applied treatments, supply details and facts that will help in the formulation of effective countermeasures to minimize deconditioning effects currently associated with exposure to actual or simulated weightlessness.

B. Background and Significance

Biochemical adaptation of enzyme activity levels of oxidative enzymes in rats are well documented with regard to exercise training (8, 27, 35), dobutamine administration (72) and suspension hypokinesia/hypodynamia (H/H) (19). Although the specific adaptive change in each of these three situations have been explored individually, there is little information currently available with regard to any interactive effects that the three treatments might have. Studies reporting alterations in antioxidant enzymes (34, 37) and myosin isoenzyme patterns (59) with exercise are available, but information relating antioxidants and myosin isoenzyme with dobutamine treatment or hindlimb suspension are virtually nonexistent.

Common adaptive changes in oxidative enzymes in rats with 8-12 weeks of exercise training include significant increases in citrate synthase (CS) and succinate dehydrogenase (SDH) (35,38). Similar increases have been demonstrated in a variety of strains, in both sexes and at different intensities of treadmill training (3).

Dobutamine, a synthetic catecholamine, has been shown to elicit cardiovascular and metabolic conditioning effects similar to those

produced by exercise training (42, 44, 72). Although there is no information available on dobutamines ability to alter oxidative enzyme levels in rats, there are reports of significant ($p \leq .10$) alterations of citrate synthase in a bed rest study with humans (72). A recent study comparing the effects of dobutamine administration to a 14 week exercise program with rats indicated that dobutamine was able to induce many of the same changes seen with the exercise training program (15).

Hindlimb suspension of rats results in alterations to the contractile and morphometric properties of certain weight-bearing muscles (22). Decreases in SDH activity with suspension have been reported in the medial gastrocnemius, the tibialis anterior and the soleus (4, 22). Myosin ATPase activity in light and dark fibers in the tibialis anterior were unaffected by suspension, while a 20% increase of dark staining fibers in the soleus was reported. These results support physiological data suggesting that there is a conversion from slow to fast fibers with suspension (4).

Chronic aerobic exercise has been shown to alter the biochemical and mechanical properties of the heart. When compared with hearts of sedentary controls, hearts of exercised rats contract faster, relax faster, and show elevated ejection fraction (66). These changes in contractile activity are associated with alterations in the biochemical properties of the myofibrils. The mechanism for change in the properties of myosin is not agreed upon, but recent evidence suggests that the increase in activity of myosin in hearts of exercised rats may be a reflection of the distribution of myosin isoenzymes (59). Myosin isoenzyme profiles indicate that endurance exercise produces a net increase in the relative amount of V1 myosin and that this shift in the isoenzyme profile is proportional to the increase in myosin Ca^2-ATPase activity of the rats (59).

Investigation into the alterations of oxidative and antioxidant enzyme activity levels and variations of myosin isoenzyme patterns in cardiac muscles in this study was of particular importance in light

of the treatments being applied. Examination of exercise treatment prior to suspension H/H and dobutamine treatment during the suspension provided a variety of interactive conditions which allowed these enzymes to be analyzed, described and compared under these particular conditions for the first time.

Metabolic control factors relating to glucose metabolism in the skeletal muscle and the blood have been studied extensively. The effects of exercise, catecholamines and situations designated to simulate weightless on a variety of metabolic control factors have been well documented in terms of individual treatments. Again, as was the case with biochemical adaptations, there is very little information available with regard to the interactive effects of the treatments on metabolic control factors.

Skeletal muscle comprises the major mass of insulin sensitive tissue in both rodents and man (28, 39, 50). Alterations in skeletal muscle insulin action can be caused by changes in either insulin sensitivity or responsiveness, or both. Changed insulin sensitivity is normally related to a change in the prevailing insulin concentration which affects the insulin receptor number causing a shift either right or left in the insulin dose response curve (30). On the other hand, a changed insulin responsiveness is caused by intracellular alterations in the routes of transport or disposal of glucose (36). When a tissue's insulin responsiveness of glucose metabolism is reduced, either the glucose transport processes impaired, or intracellular enzyme function is depressed. In normal tissue, insulin activities glycogen synthase and the glucose taken up by the cell is converted to glycogen (30). In insulin unresponsive tissue, this route of glucose disposal is impaired and excessive lactate is produced (32). Additionally, glycogen levels are elevated and glycogen synthase activity is depressed (32). An indirect method to determine if insulin resistance exists in a tissue which cannot be assayed directly for insulin responsiveness (as in the tissue utilized in this study), is to assess the "normality" of one of the tissue's pathways for the disposal of glucose: hexokinase and glycogen synthase, and the glycogen content.

Exposure of rats to prolonged weightlessness (58) or hindlimb hypokinesia (70) causes a substantial loss of body weight and skeletal muscle mass. Because reductions in circulating somatomedins (insulin-like growth factors) are associated with a decrease in growth (67) and increased somatomedins can enhance muscle (52) and total bodily growth (26, 67), it seems reasonable to hypothesize that reductions in circulating somatomedins or in their growth-promoting activity might play a role in hypokinesia-induced loss of body weight and muscle mass.

Because regular physical exercise either maintains or increases muscular growth, it might be expected that exercise would increase somatomedin concentrations or activity. If this were so, because somatomedins have a relatively long half life, i.e., 3 h in hypophysectomized rats (11) and 20 h or more in man (5), the increased somatomedin activity might have a residual effect for many hours following the exercise. However, there is at present no clearly established association between exercise and somatomedins. Chronic physical exercise in humans, when prescribed in a manner designed to cause a net caloric deficit, has been shown to decrease concentrations of somatomedin-C in serum (71). On the other hand, a strenuous 9-week weight training regimen increased plasma concentrations of somatomedin-C by 35% in adult normal men who consumed a high energy diet (2). Furthermore, in hypophysectomized rats, extracts of hypertrophied skeletal muscle exhibited elevated somatomedin-like activity in a chick embryo fibroblast bioassay relative to contralateral control tissues (71). Thus, there is some support based on previous work for the idea that exercise might minimize the somatic atrophy that accompanies suspension hypokinesia by increasing somatomedin-C concentrations and/or growth promoting activity.

Cardiovascular response and control are included among the many biomedical areas that are affected by spaceflight. The syndrome of cardiovascular instability that results has sometimes been termed "cardiovascular deconditioning" and the manifestations associated

with the syndrome are most apparent when upright posture is assumed after re-entry. Primary symptoms include tachycardia, hypotension, reduced exercise capacity and labile blood pressure (43). Cephalad redistribution of fluid caused by the lack of hydrostatic pressure is often cited as the probable cause of most of the deconditioning response. Although the exact biologic mechanisms of the deconditioning syndrome have not been established, a number of hypotheses have been formulated. Two proposed hypotheses that have direct bearing on this study are 1) possible changes in reflex control of cardiovascular function and 2) impaired release or uptake of norepinephrine.

While suspension H/H in rats is currently being used as an animal model for simulating weightlessness, there are very few studies that have investigated the adaptive cardiovascular response with this model (68). Two studies (60, 68) that have examined cardiovascular changes with head-down suspension, while differing in some respects, provide evidence that early central venous changes and cephalad fluid shifts that are characteristic of weightlessness can be produced in rats subjected to this treatment. Discrepancies in results between the two studies appear to be based on the difference in methods of suspension. Shellock, et al., (68) utilized tail-traction method that is very similar to the one employed in this study, while Popovic et al., (60), used a harness technique. Later results reported by Popovic indicated that some of the inconsistencies between his earlier results and reported hemodynamic response to head-down tilt in man may have been due to elevated levels of stress-associated hormones (60). It was suggested that stress from the animals struggling to escape the harness may have inadvertently affected the cardiovascular response to simulated weightlessness. So, although there are indications that tail-traction suspension of rats may be a more appropriate model for the study of cardiovascular change in simulated weightlessness than some other methods, it is also clear that more work needs to be done to characterize the model.

Cardiovascular responses to dynamic exercise have been well documented. The sympathoadrenal system plays an integral function in mediating acute responses to dynamic studies. A number of studies have shown increased levels of adrenergic amines in the blood during dynamic exercise (1). Recent work has also demonstrated that the administration of catecholamines and cardioactive sympathomimetic agents, such as dobutamine, are associated with increased cardiac function similar to that observed during dynamic exercise. Chronic infusion or injection of dobutamine has also been reported to produce cardiovascular effects similar to those occurring after exercise training (44).

In addition to sympathetic activity, diminished input from the parasympathetic nervous system is also part of the normal responses to exercise. Alterations in heart rate response is determined by a number of factors including balance between sympathetic and parasympathetic efferent activity as well as by humoral catecholamines (21). Continued investigation into the reflex control of heart rate and the relative roles of the sympathetic and parasympathetic systems with regard to the three treatments examined in this study should provide valuable information relating to mechanisms of control that are affected by space flight.

C. Design and Methods

The basic experimental design of the primary experiment (Phases II-III) is described in detail in the attached "Study Protocol" (Appendix A). Briefly, the design is a randomized 2^3 factorial design with approximately 6 observations in each treatment combination. This design allows for statistical evaluation of the main effects of each of the treatment factors (training status, drug status and suspension status), as well as their interaction. Most of the analyses included in this report were carried out on tissues from the 48 animals described in Phase II and III of the protocol. A schematic showing the various subgroups and the n of each follows.

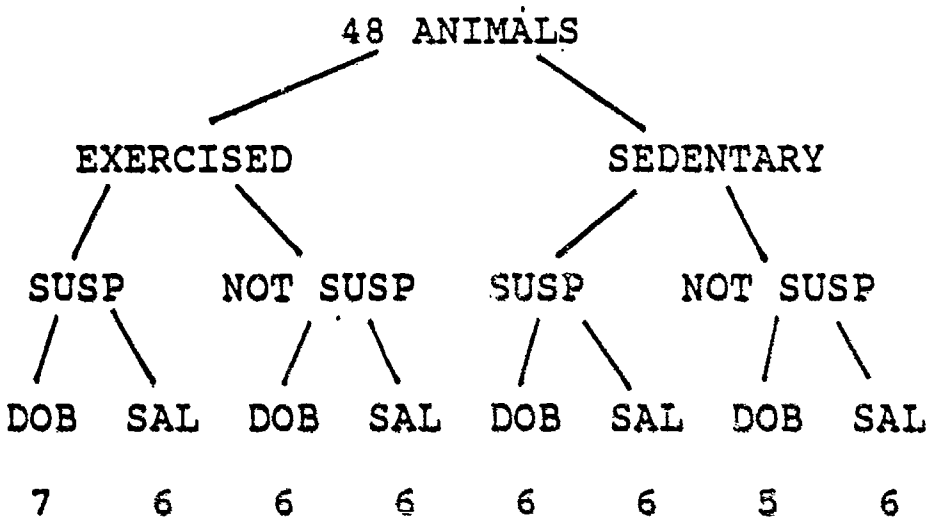


Figure 1.

Additional animals were utilized to begin development of methods that would allow study of the arterial baroreflex control of heart rate and blood pressure and the relationships between the control mechanisms and the treatments. These preliminary studies were conducted in the laboratory of Dr. R. L. Hamlin in the College of Veterinary Medicine, The Ohio State University. A total of 24 animals were used for this additional phase of the investigation. Procedures and methods that were utilized during this part of the experiment are described briefly in the cardiovascular methods section below, and in more detail in Appendix D.

Specific methods for the three major areas of the investigation are described below. The methods described were those used during the analysis and in a few instances differed slightly from those described in the proposal.

1. Enzyme Adaptations: These analyses were done in Dr. A.J. Merola's laboratory in the Department of Physiological Chemistry, College of Medicine, The Ohio State University. Specific areas investigated were 1) oxidative training enzyme levels of

citrate synthase (in the soleus), 2) antioxidant enzyme levels in the soleus (superoxide dismutase and catalase), and 3) myosin isoenzyme profiles in the left ventricle.

Citrate synthase was measured by the spectrophotometric method of Saere (62). The reaction involves the measurement of citrate synthetase activity by linking the release of CoASH to the colormetric agent DTNB 5, 5-dithiobis - 2-nitrobenzoate. Changes in absorbence are followed at 412 nm. Calculation of units of activity involves the use of the extinction coefficient 13.6.

Superoxide dismutase was measured as described by Keele, et al. (40) by its enzymatic ability to inhibit the O_2 reduction of ferricytochrome C. Tissue samples were homogenized in iced 0.05M K phosphate buffer, pH 7.8 containing 100 uM EDTA. The crude homogenates were centrifuged at low (700 xg) speeds and the supernatant assayed with proper dilution for SOD activity reaction mixtures containing 1 x 10^{-5}M ferricytochrome C, 5 x 10^{-5}M xanthine, xanthine oxidase, and 1 x 10^{-5}KCN. Under these conditions 1 unit activity is that amount of enzyme which halves the rate of reduction of cytochrome C when the base rate is 0.015 A/Min.

Catalase was measured by the method of Cohen, et al., (11) by titrating residual H_2O_2 remaining after a known time with permanganate. The reaction mixture consists of 0.01M phosphate, pH 7.0, 6mM H_2O_2 and 0.5 ml homogenate in a fixed volume of 5.5 ml. After three minutes, the reaction mixture was acidified with 1.0 ml H_2SO_4 and the remaining H_2O_2 measured addition of 7 ml 0.01 KMNO4. Residual KMNO4 was measured at 480 nm and the activity was reported as the first order rate constant k per g wet weight.

Myosin isoenzyme profiles were measured by pyrophosphate gel electrophoresis according to the method of Hoh et al., (25). The

electrophoresis was performed at 88V (14V/cm, constant voltage) for 19 hours at 2°C. Gels were stained overnight in 0.025% Coomassie brilliant blue R-250, 25% isopropanol, and 10% acetic acid, then destained in 10% acetic acid. The destained gels were scanned with a high resolution laser densitometer.

2. Metabolic Control Factors: These analyses were done in the Exercise Physiology Laboratory at The Ohio State University. Drs. Sherman and Lamb oversaw these analyses. Specific areas investigated were 1) somatomedin-C in plasma and muscle, and 2) glycogen and related substances (glycogen, glycogen synthase, hexokinase and protein).

Plasma somatomedin-C was analyzed by radioimmunoassay (RIA) procedures utilizing a competitive binding protein assay with a double antibody reaction (Nichols Institute). A recovery of 81.4 ± 2% (SE) of known KSI-labelled standards added to the plasma samples was achieved.

Muscle somatomedin-C from the soleus and plantaris muscles was dissociated from binding proteins by an acid extraction prior to radioimmunoassay. To obtain adequate volumes of tissue, frozen soleus and plantaris muscles of one hindlimb of each rat were pulverized with a mortor and pestle under liquid nitrogen. Five volumes of 1 M acetic acid were added to 0.05-0.5 g of the combined tissue in a polypropylene centrifuge tube. The samples were vigorously shaken for 15 s, incubated on ice for 2 h, and centrifuged at 4°C for 10 min at 1000 x g. The supernatant liquid was neutralized with 5 M sodium hydroxide, and somatomedin-C was measured in duplicate by radioimmunoassay as described for plasma.

Glycogen synthase activity was measured without added sulfate at several concentrations of G6P: 0, 150, 500, 100, 1500, and 10000 uM. The procedure was that of Kochan et al., (41) and is a modification of the method described by Thomas et al., (73). An

aliquot of the tissue extract was diluted 5-fold (vol/vol) with 50 mM morpholinopropane sulfonic acid (MOPS) buffer, 25 mM KF and 20 mM EDTA (pH 6.9 at 25°C). Glycogen synthase activity was then determined by adding 0.04 ml of the tissue extract to 0108 ml of reaction cocktail containing 0.045mM [u-^{14}C]-uridine diphosphate glucose (UDPG) (5000 cpm/nmol), 10.6 mg rabbit glycogen/ml, 15mM organic phosphate, 50 mM MOPS buffer, 25 mMKF, 20 mM EDTA and the appropriate concentration of G6P (pH 6.9 at 25°C). Glycogen synthase total activity was expressed as umol/min/mg protein. The glycogen synthase activity ratio was calculated by dividing the activity in the presence of 0 uM G6P by the total activity (10,000 uM G6P). The glycogen synthase fractional velocity was calculated by dividing the activity of glycogen synthase at the subsaturating concentration of G6P (500 uM G6P) by the activity of glycogen synthase at the maximal G6P concentration (10,000 uM).

Hexokinase in the tissue extract was determined fluorometrically by following the reduction of NADP+ using glucose as substrate (45). An aliquot of tissues extract was added to 1 ml reaction cocktail containing 100 mM tris (hydroxymethyl) - aminomethane hydrocholoride (TRIS-HCl) buffer, 1.0 mM glucose, 2 mM adenosine triphosphate (ATP), 5 mM magnesium chloride ($MgCl_2$), 0.1 mM NADP+, 0.02% BSA, and 4.5 u G6P dehygrogenase/ml (pH 8.1 at 25%C). The hexokinase activity was expressed as umol/min/mgpr.

Muscle glycogen was determined after acid hydrolysis (2N HCl) of a weighted sample of muscle (20 mg) at 90°C for 2 hrs. After neutralization (NaOH), an aliquot of the glycogen hydrolyzate was added to 1.0 ml of a reaction cocktail containing 25 mM TRIS-HCL, 1.0 mM MG Cl_2 0.5 mM dithiothreitol (CTT), 0.3 mM ATP, 6.05 mM NADP+, 0.02 u G6P/ml, and 1.0 u hexokinase/ml (pH 8.1 at 25°C) (36). The concentration of hydrolyzed glucose residues was expressed as umol/mg protein.

Protein concentration was measured using Commassie-blue dye binding procedure after digestion in 0.01 M NaOH (Bio-Rad). Bovine serum albumin was used as standard.

3. Cardiovascular Response and Control: These studies were carried out in the laboratory of Dr. R. L. Hamlin in the College of Veterinary Medicine, The Ohio State University. Specific areas investigated were heart rate (HR) and electrocardiographic (ECG) waveform response, 2) preliminary investigation into methods to study arterial baroreflex control of heart rate and blood pressure (BP) and 3) histological determination of cardiac fibrosis and left ventricular (LV) wall area.

HR and ECG response was examined after the animal had been anesthetized with an IP injection of ketamine/xylaxine. Electrode placement allows for simultaneous recording of leads I, AVF and V_{10} on a Honeywell E for M model photographic oscillograph. Heart rate and total peak to peak QRS amplitude in lead AVF were calculated from the ECG tracing. Additional measurements of interest included P wave in lead I, R wave in I, and the S/R ratio in lead AVF.

Baroreflex control of HR and BP was first determined by measurement of arterial baroreflex sensitivities or the HR/ mean arterial pressure (MAP) and is expressed as beats $min^{-1}mm$ of Hg^{-1}. Measurement of aortic pressure was made by advancing a polyethylene catheter to the aortic root after its percutaneous introduction into the subcutaneous segment of the right common carotid artery. Once placed, the cannula was exteriorized behind the neck, flushed with heparinized saline, and plugged with a sealed 23-gauge needle. A calibrated pressure transducer was connected to the cannula with a zero pressure reference at the point of the shoulder. Pressure tracings and ECG waveforms were displayed and recorded on a Honeywell E for M model photographic oscillograph. Alternate cannulation methods were also utilized and are described in more detail in Appendix D.

Histological heart sections were obtained by removing a 2 mm slice of myocardium from the apex of the heart along a line parallel to the atrioventricular grove. A second 2 mm section is obtained and fixed in 10% buffered formalin. The fixed section of the left ventricle was placed in a tissue cassette, dehydrated and embedded in paraffin (46). Five um sections were cut from the slice, then mounted and stained with Masson's trichome to allow for microscopic determination of fibrosis. A Zeiss Interactive Digital Analysis System (ZIDAS) set in the delta area program was used to measure left ventricular areas. This was accomplished by tracing the cursor along the epicardial surface of the left ventricle and septum in a clockwise manner. The area (mm^2) from this measurement was stored in the active channel. Then the cursor was used to trace the endocardial surface of the left ventricle in a counterclockwise fashion. This area was then substracted from the area data in the active channel. A schematic representation of this procedure is shown in Figure 1.

D. Results and Discussion

Information obtained for both the exercised and sedentary groups are combined and presented together in this section. Previous results that appeared in the proposal are provided in Appendix B. Results presented here are discussed with regard to relationships between different variables as well as with information available from pertinent literature. Before presenting results from the three major areas of investigation, the basic group characteristics of body weight and food consumption are described below.

Food consumption and body weights for the eight groups of rats are shown in Figures 2-5. The basic patterns of food consumption between the four exercised groups and four sedentary groups are very similar with the largest decreases in food consumption over the first three days of treatment. The suspended groups consumed less food than the control animals with the sedentary groups eating 7-9 grams

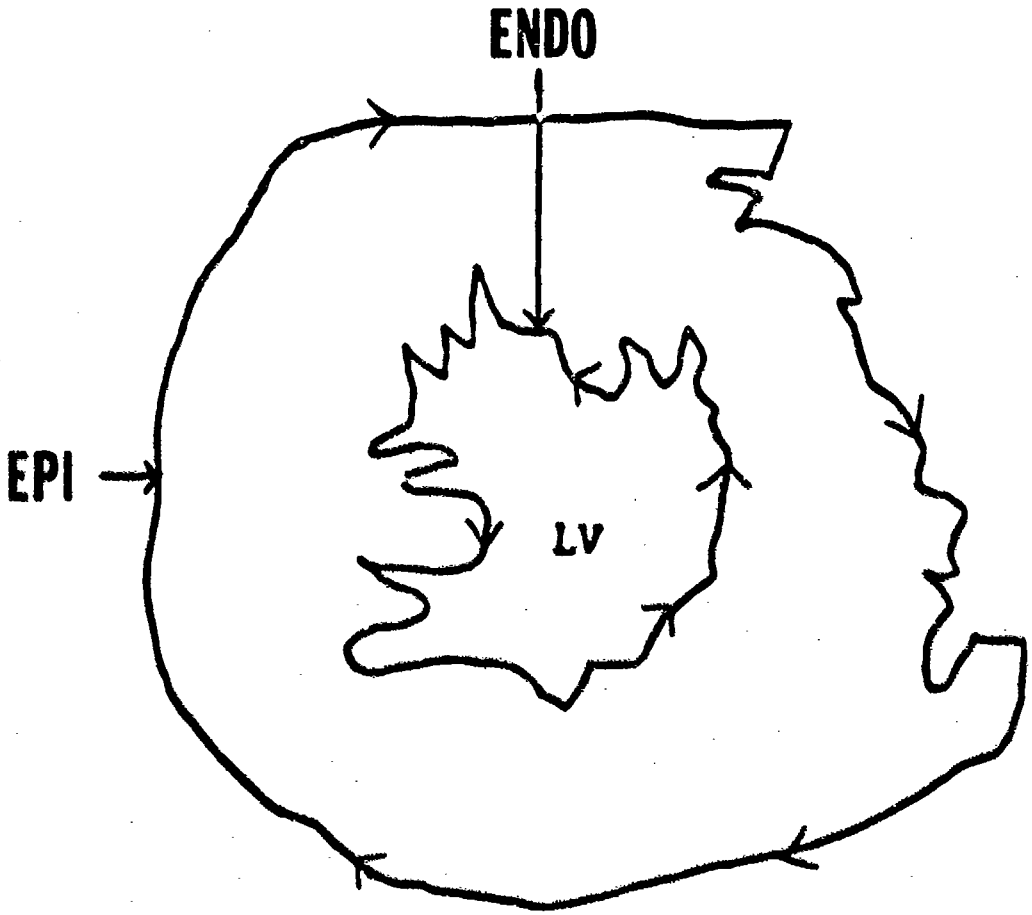


Figure 1.

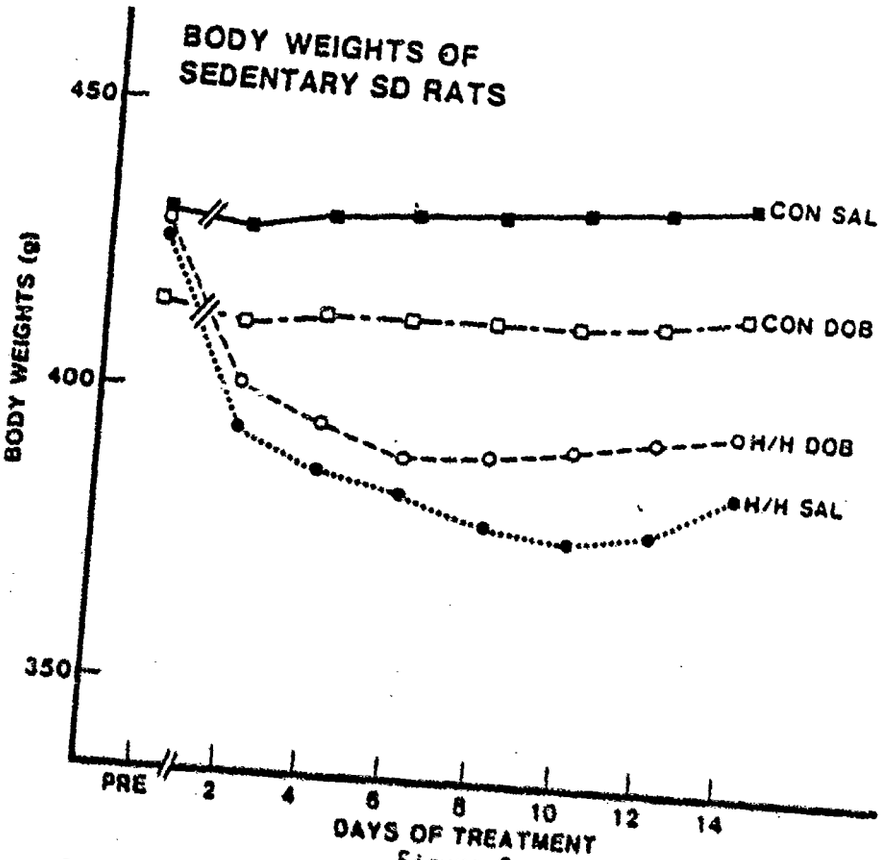


Figure 2.

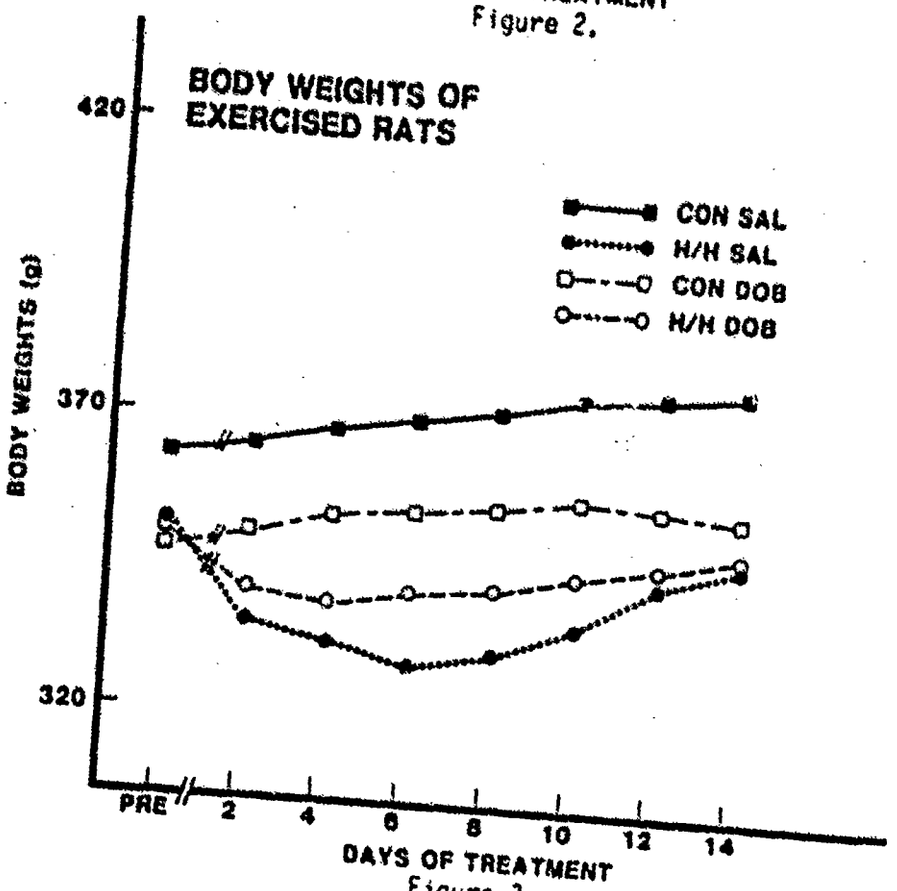


Figure 3.

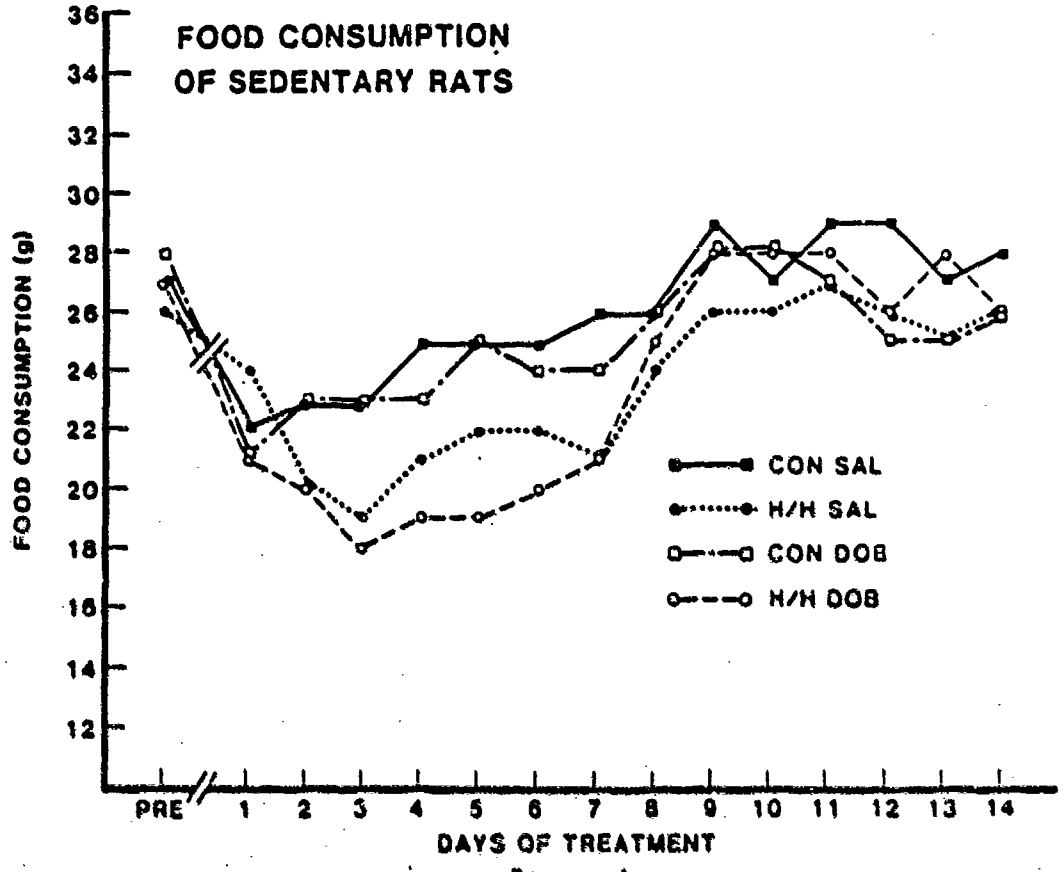


Figure 4.

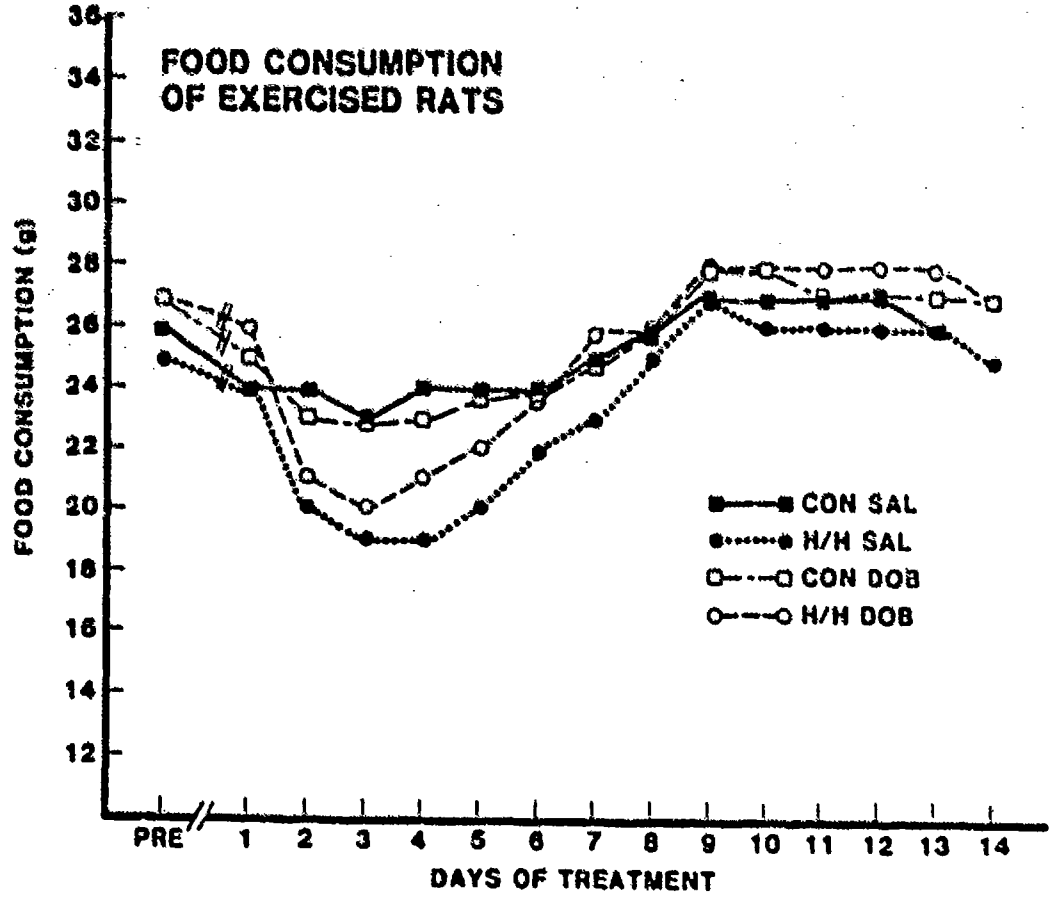


Figure 5.

less than the pretreatment values and both exercise groups consuming approximately 6 grams less than their pretreatment values. All eight groups started to increase food consumption after three days and all were back to baseline values by seven days. The changes in body weight were again similar in general patterns, but it is important to note that the maximal percent change from pretreatment weights were somewhat different. The suspended groups of sedentary animals lost a larger maximal percent body weight (~11% without dobutamine and~9% with dobutamine) than did the exercise animals (~ 6% without dobutamine and~3% with dobutamine).

1. Enzyme Adaptations: These analyses were conducted in Dr. Merola's laboratory, Department of Physiological Chemistry, College of Medicine. Dr. Girten completed the assays and supplied the results reported in this section. Dr. Merola was available for advice and support throughout each of these procedures. A three way ANOVA was done for each variable and there was significant ($p \leq .05$) interaction between suspension and dobutamine treatments for the enzymes measured. Tukey's pairwise comparisons were used to determine significant difference between pairs of groups.

   Citrate synthase is an oxidative training enzyme that has been shown to increase with aerobic training (62), decrease with hypolinesia/hypodynamia suspension (19), and increase with chronic intravenous dobutamine infusion (72). The citrate synthase levels for the eight experimental groups are shown in Figures 6 and 7. The intra-assay and inter-assay coefficients of variation for the assay were 3.9% and 2.2%, respectively.

   Overall ANOVA indicated significant differences and Tukey's showed that the four exercise groups had greater levels ($p \leq .05$) of citrate synthase when compared with their sedentary group counterparts. The suspended groups levels were all lower than the corresponding control groups, but the differences in the animals receiving saline were significantly ($p \leq .05$) lower while

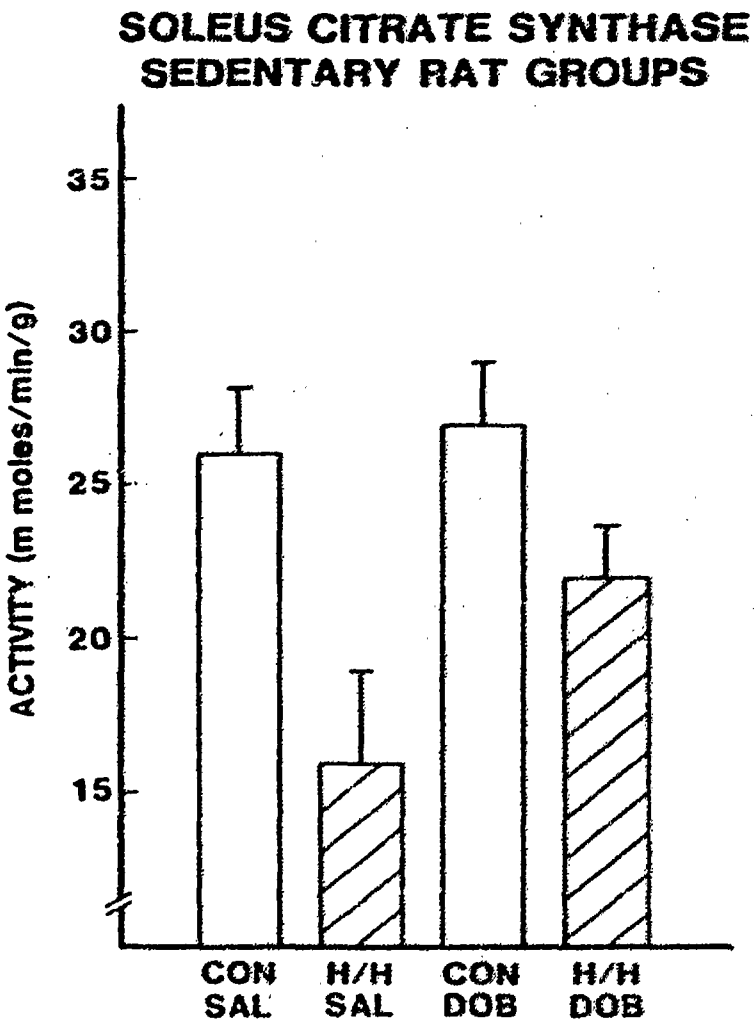


Figure 6.

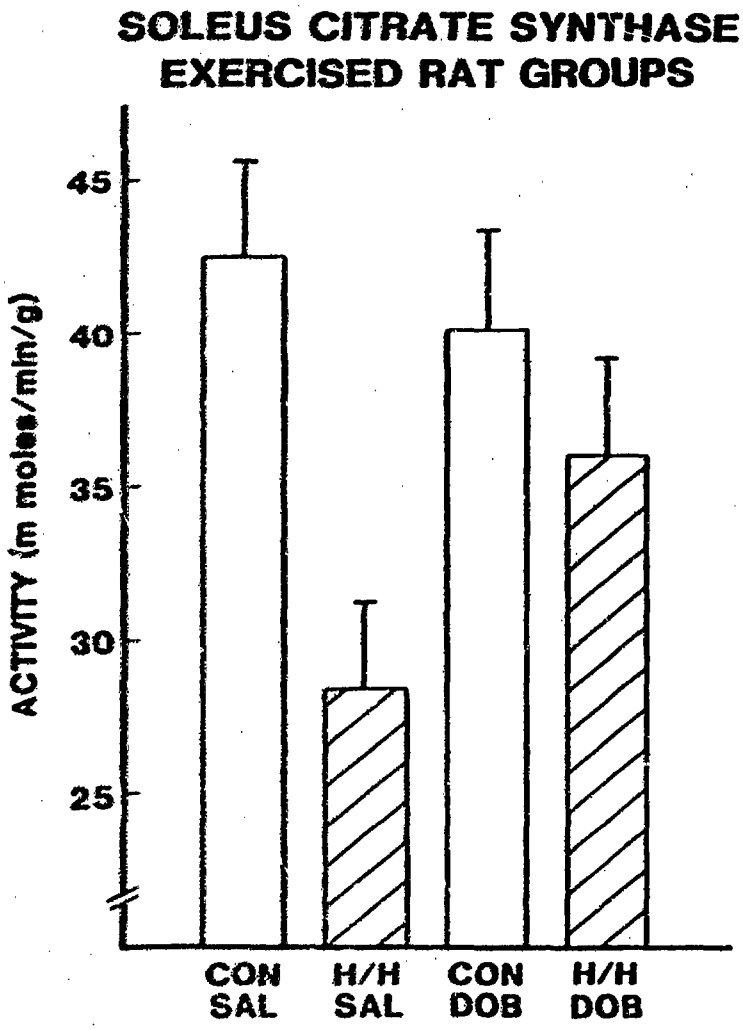


Figure 7.

those receiving dobutamine were not. These results suggest that although exercise trained animals appear to lose more citrate synthase from suspension as compared to the sedentary rats, the decrease expressed as percent of control levels is somewhat lower (~32% vs 38%) and the final value is significantly greater than the suspended animals that were not exercised trained. The greater levels of citrate synthase seen with dobutamine appears to have almilerated the negative effect of suspension.

These findings were in agreement with our original hypotheses that exercise training would result in greater citrate synthase values and that dobutamine might inhibit the negative effects that suspension hypokinesia/hypodynamia has on oxidative training enzyme values. Values from the present study correspond well with control and suspended animals from previous studies (19, 62). Although positive effects ($p \leq .10$) of dobutamine on citrate synthase have been reported in a previous bed rest study (72), the positive results reported here are more definitive ($p \leq .05$). This may, in part, be due to better experimental control and subject homogeneity in animal studies when compared to human studies, but might also be related to length of experiment and tissue sampling procedures.

Catalase is an antioxidant enzyme that is predominately found in the peroxisomes of cells and often works in concert with glutathione peroxidase to keep hydrogen peroxide at a low steady state level (56). Catalase levels from the soleus muscle are shown in Figures 8 and 9. The intra-assay and inter-assay coefficients of variation for the assay were 3.7% and 2.4%, respectively.

Statistical analysis indicated significant group differences and post hoc comparisons showed that three of the four exercise groups had significantly ($p \leq 0.05$) greater levels of catalase, than their sedentary counterparts; the only exception was the suspended saline groups ($p = .06$). Catalase levels in the

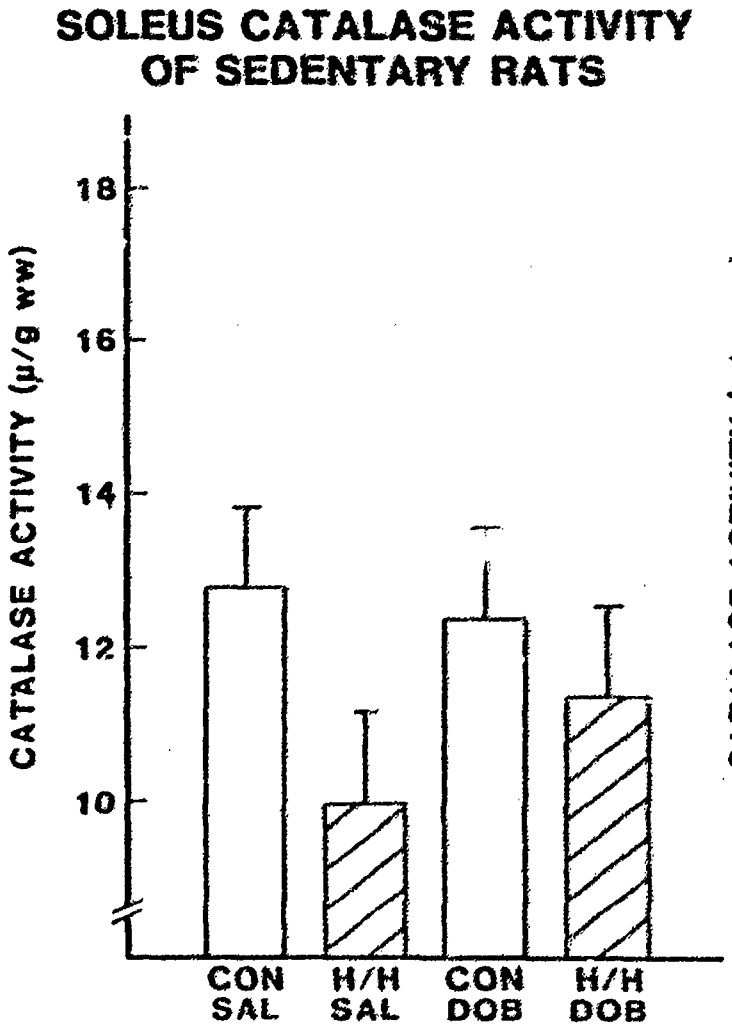


Figure 8.

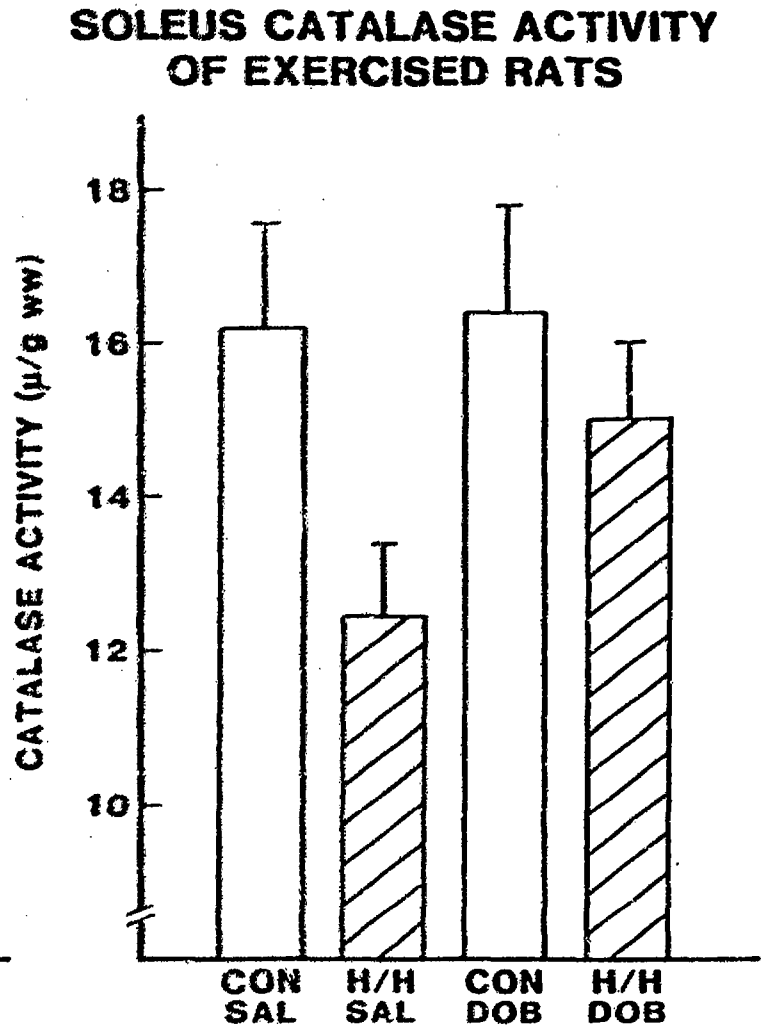


Figure 9.

suspended saline groups were significantly ($p \leq .05$) lower than those found in the control saline groups.

Although catalase has not been examined with regard to weightlessness or simulated weightlessness it has been reported to correlate ($r=.72$) well with muscle oxygen consumption, which can serve as an oxidative training marker in a manner similar to citrate synthase. The correlation between catalase and citrate synthase in the present study ($r=.74$) supports results reported by Jenkins et al., (35). These findings indicate that exercise training prior to simulated weightlessness and dobutamine administration during the simulations may be helpful to prevent large decreases in muscle catalase levels that result during this time.

Superoxide dismutase is an antioxidant enzyme that catalyzes the destruction of the superoxide radical and has been shown to correlate ($r=.70$) well with catalase (35). The superoxide dismutase levels for the eight experimental groups are shown in Figures 10 and 11. The intra-assay and inter-assay coefficients of variation for the assay were 4.1% and 1.9%, respectively.

Overall ANOVA indicated significant differences and post hoc analysis showed that each of the exercise groups had greater levels ($p \leq .05$) of superoxide dismutase when compared to their sedentary counterparts. The suspension treatment in the saline groups resulted in significantly ($p \leq .05$) lower enzyme levels when compared to the control saline groups.

The eight experimental groups display a very similar relationship to each other when comparing superoxide dismutase, catalase and citrate synthase results. This finding and the data obtained are in general agreement with previously reported information, but it is important to note that antioxidant enzyme activity has not been examined with relationship to suspension treatment or with regard to the various interactions that were

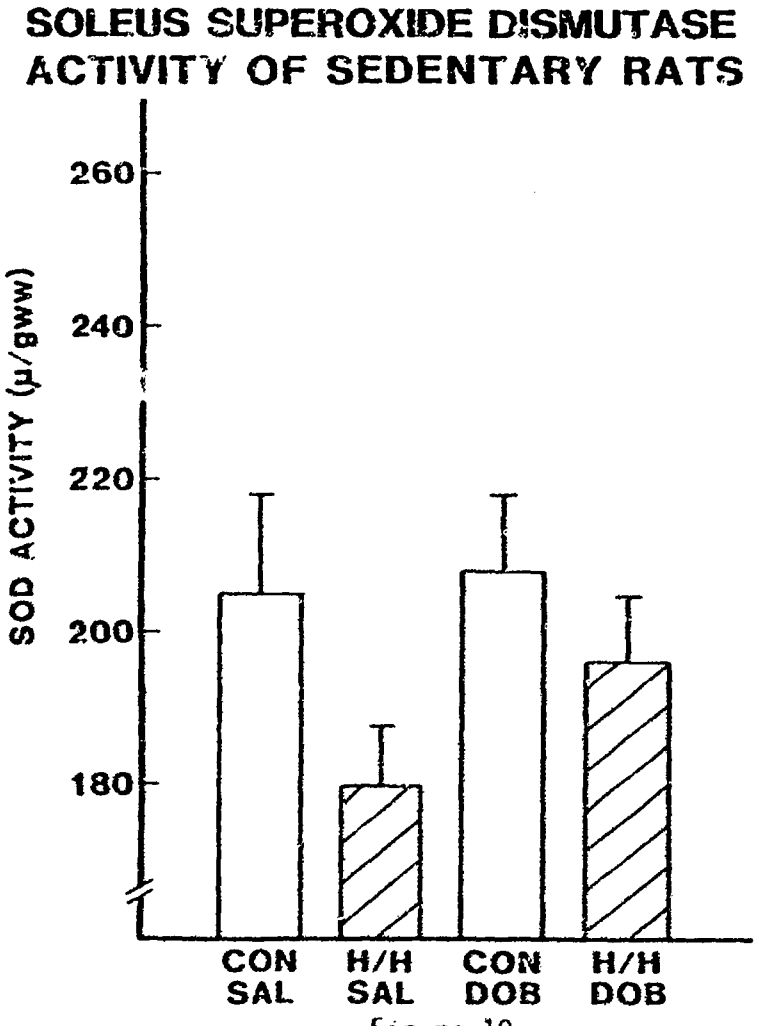


Figure 10.

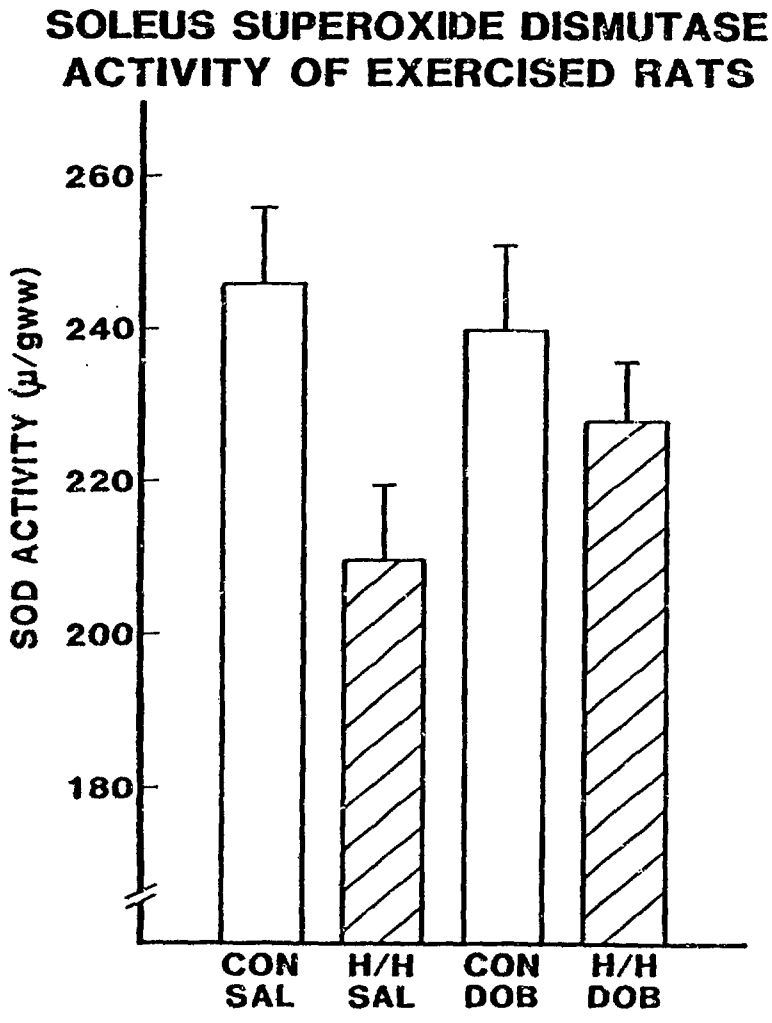


Figure 11.

created by the design of the current study. With this thought in mind then, it is of particular importance to emphasis that suspension (and perhaps other simulations of weightlessness) not only results in decreased oxidative enzyme levels but also results in decreased levels of two antioxidant enzymes that are normally necessary to prevent uncontrolled free radical reactions (56) that can lead to oxidative tissue damage.

Myosin isoenzyme profiles from the ventricles of rat hearts reflect the contractile activity of the myocardium and are associated with the biochemical properties of the myofibrils. Representative samples of the ventricular myosin isoenzyme profiles for each of the eight experimental groups are shown in Figures 12 and 13. The mean $\pm$ standard deviation for the groups are summarized in Table 1.

The ANOVA comparison indicated significant group differences and Tukey's pairwise comparisons indicated that the V1 component of the myosin isoenzyme was significantly ($p \leq .05$) greater in each of the control groups when compared to their suspended counterparts and that the exercised groups had significantly ($p \leq .05$) greater V1 values than their sedentary counterparts. Groups that received dobutamine injections and were suspended had significantly ($p \leq .05$) greater V1 isoenzymes than did their saline group counterparts, and although the control dobutamine groups had slightly larger V1 components than did the control saline groups, this difference was not significant. There were no significant differences in the V2 components among the four control groups and the only significant difference ($p \leq .05$) in the suspended groups occurred between the sedentary dobutamine group and the exercise dobutamine group with the sedentary groups V2 being greater. The V3 component of the sedentary groups are in each case significantly ($p \leq .05$) greater than their exercised counterparts and the only instance that the V3 isoenzyme is significantly ($p \leq .05$) different when comparing saline to dobutamine occurs in the suspended sedentary animals.

REPRESENTATIVE MYOSIN ISOZYME PROFILES
FROM CONTROL GROUPS OF RATS

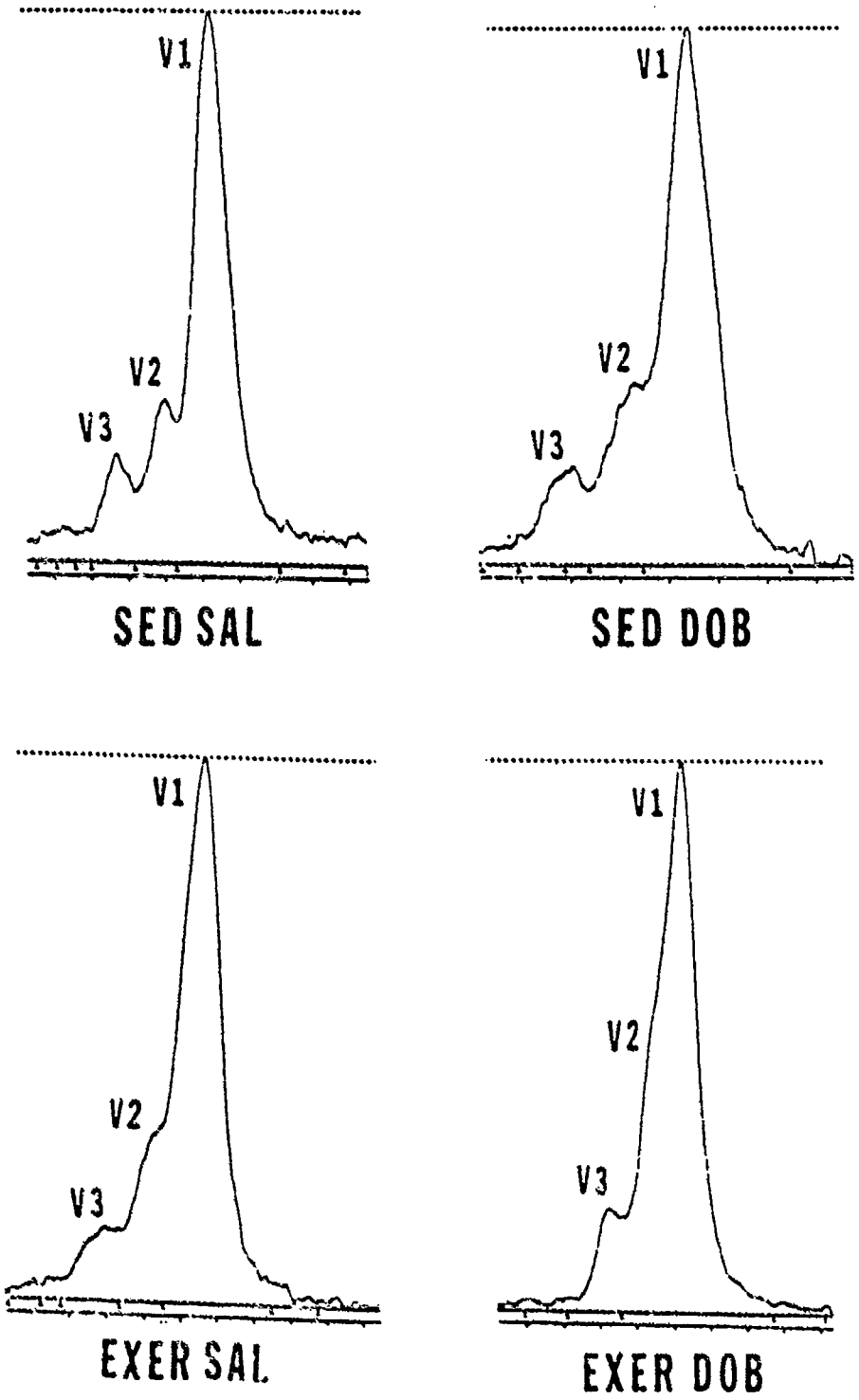


Figure 12.

REPRESENTATIVE MYOSIN ISOZYME PROFILES
FROM SUSPENDED GROUPS OF RATS

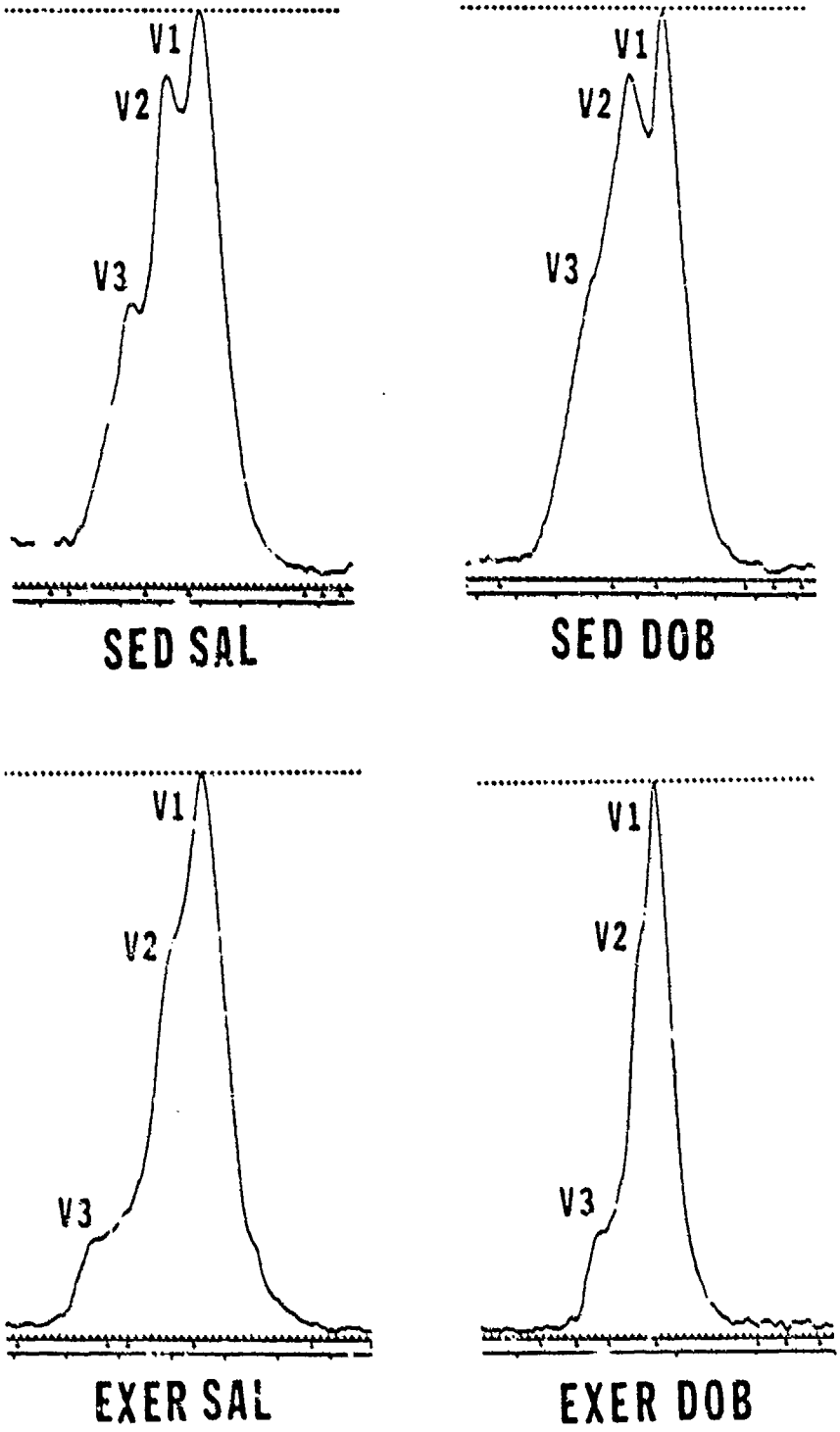


Figure 13.

Table 1

Percent Area of Rat Ventricular Myosin Isoenzyme Components

	Control Groups		Suspended Groups	
	Sedentary Saline	Sedentary Dobutamine	Sedentary Saline	Sedentary Dobutamine
V1	70.6 ± 4.2	73.6 ± 4.0	30.0 ± 3.1	47.3 ± 4.4
V2	13.4 ± 2.1	12.3 ± 2.8	35.6 ± 4.7	33.7 ± 5.0
V3	16.0 ± 3.0	14.1 ± 2.0	26.0 ± 3.8	19.0 ± 3.6
	Exercise Saline	Exercise Dobutamine	Exercise Saline	Exercise Dobutamine
V1	50.1 ± 3.8	83.9 ± 4.1	58.8 ± 3.4	10.7 ± 4.6
V2	10.9 ± 2.61	7.3 ± 2.0	27.2 ± 4.1	21.0 ± 3.2
V3	9.0 ± 2.9	8.8 ± 2.7	14.0 ± 3.6	8.3 ± 5.9

The shift from the V2 and V3 isoenzymes to the larger V1 component has been associated with aerobic exercise training in previous reports (59, 64) but similar shifts seen in the interactive effect of suspension and exercise and suspension, exercise and dobutamine have not been reported. In fact, the effect of suspension or other forms of simulated weightless on ventricular myosin isoenzyme patterns have not, to the best of our knowledge, been previously investigated. Therefore, the large increase in V2 seen with suspension is reported here for the first time. A profile similar to the ones reported here for suspended groups were reported for hypertensive rats (64) and this pattern was significantly improved with a 10 week swimming program. In view of this information it could be possible that the shift in isoenzyme profiles observed in the suspended rats may be due to a suspension induced hypertension, however, since blood pressures were not measured this relationship could not be tested.

2. Metabolic Control Factors: These analyses were conducted in the Exercise Physiology Laboratory. Many of the somatomedin-C assays were completed by P. Plato with help from S. Bloomfield and J. Schlub. Glycogen and related factors were done by Dr. Sherman, A. Doyle, and A. Dernbach with assistance from N. Rong and J. Goa. Dr. Lamb supplied the results reported for the somatomedin-C and Dr. Sherman supplied the results reported for the glycogen and related factors.

Somatomedin-C concentrations in plasma are shown in Table 2 and Figures 14 and 15. These results are also expressed per unit of body weight in Table 3 and Figures 16 and 17. When body weights are ignored the only significant main effect detected by the ANOVA was that for hindlimb suspension ($p<0.001$), which reduced plasma concentrations of somatomedin-C by 25% when averaged over all exercise and drug conditions. However, there was an interaction effect ($p<0.001$) between suspension and exercise treatment such that the effect of hindlimb suspension was only apparent within the subgroup of trained animals.

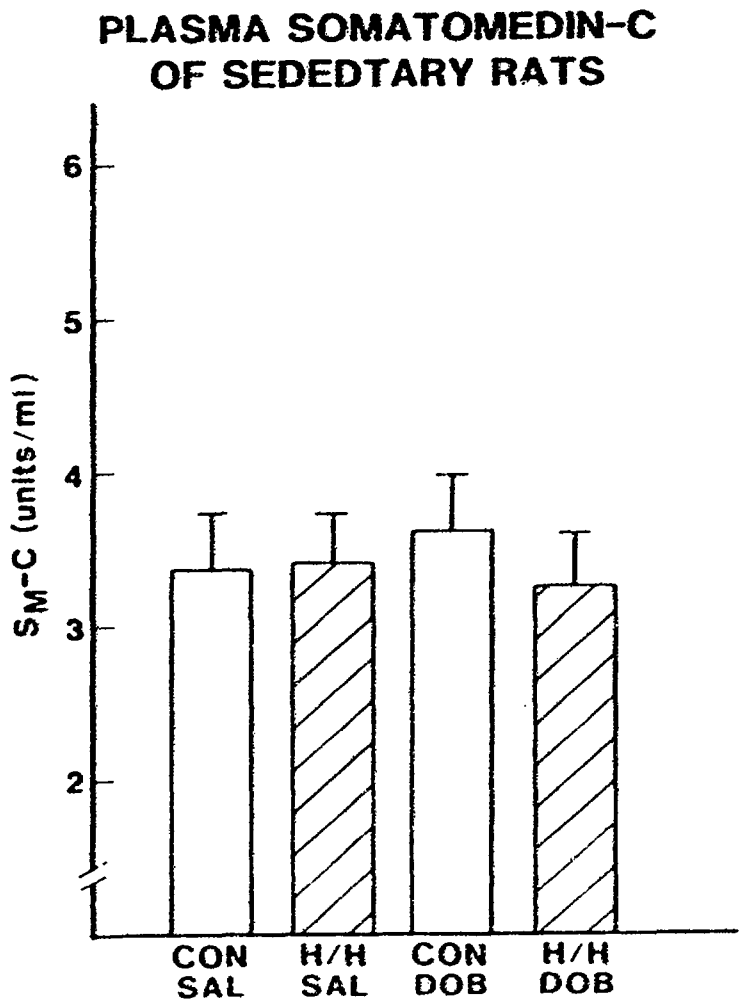


Figure 14.

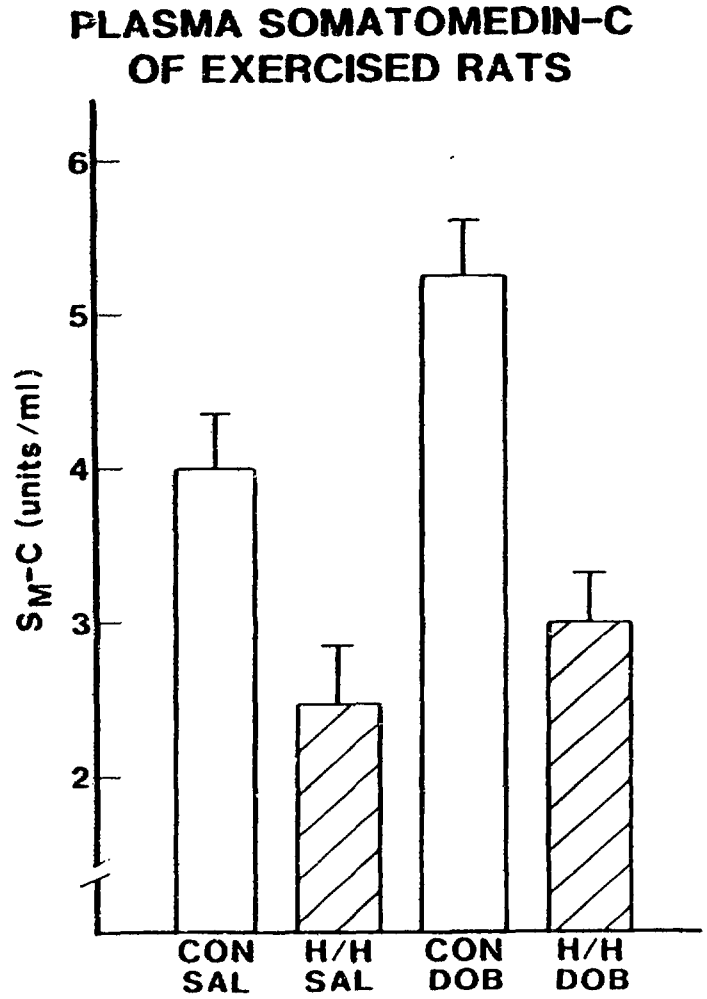


Figure 15.

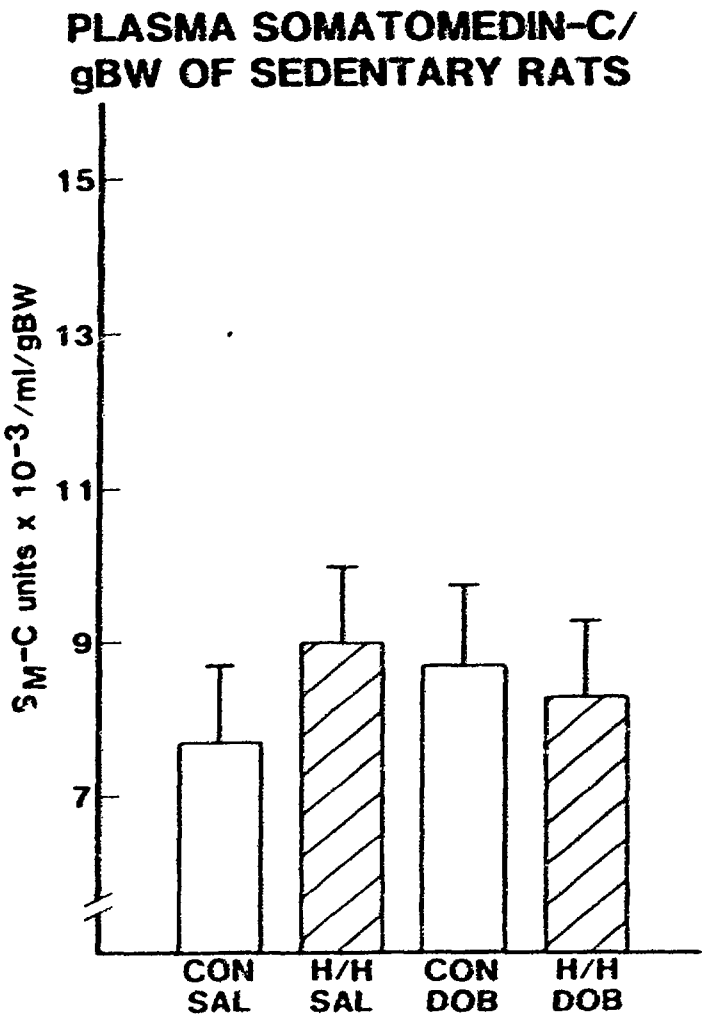


Figure 16.

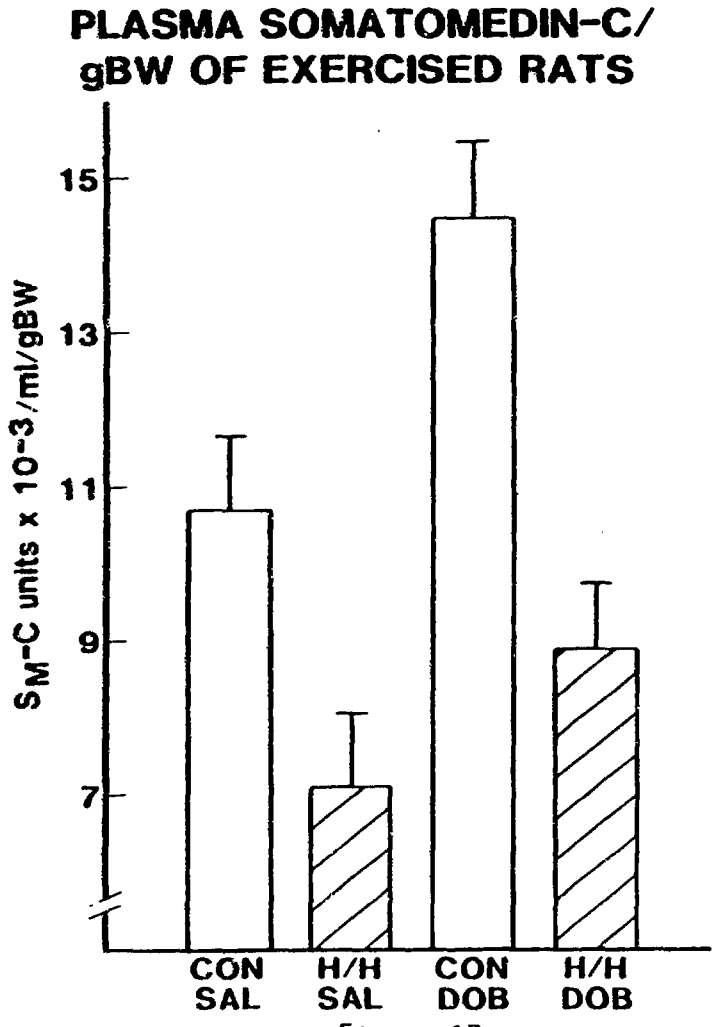


Figure 17.

Similar results were obtained when somatomedin-C results were expressed per unit of body weight (Table 3), but the ANOVA in this case also detected a significant ($P<0.01$) main effect of exercise treatment as well as a significant ($P\leq0.05$) main effect of dobutamine treatment. Once again, however, a significant ($P<0.001$) interaction between exercise treatment and suspension condition demonstrated that the exercise effect was present only in the nonsuspended rats.

Table 2

Somatomedin-C in Plasma (Means ± SE; U/mL)

	Sedentary		Exercise	
	Saline	Dobutamine	Saline	Dobutamine
Control	3.36 ± 0.35	3.60 ± 0.39$^{\$}$	4.01 ± 0.36*	5.25 ± 0.36^{+}
Suspended	3.41 ± 0.36	3.28 ± 0.36	2.46 ± 0.36#	3.04 ± 0.33

#(P<0.05) vs. $ and *
+(P<0.01) vs. all tabled values

Table 3

Somatomedin-C in Plasma (Means ± SE; mU/mL x g body weight)

	Sedentary		Exercise	
	Saline	Dobutamine	Saline	Dobutamine
Control	7.70 ± 0.98$^{\$}$	8.64 ± 1.08	10.73 ± 0.98*	14.46 ± 0.98^{+}
Suspended	8.97 ± 0.98	8.27 ± 0.98	7.05 ± 0.98#	8.86 ± 0.91

*(P<0.05) vs. $ and #
+(P<0.01) vs. all tabled values

Hindlimb suspension significantly reduced somatomedin-C values only for the trained animals. This negative effect of suspension on somatomedin-C would probably have been insignificant were it not for the fact that exercise training dramatically increased circulating somatomedin-C in the nonsuspended rats. This increase in somatomedin-C in trained rats who consumed food ad lib may correspond to the increased plasma somatomedin-C observed in young adult men who gained roughly 3.5 kg during a 9 week intensive weight training program (2), and is in contrast to the negative effects of exercise on somatomedin-C when subjects are undergoing a caloric deficit.

Contrary to the results in plasma, somatomedin-C concentrations in muscle were significantly greater in the suspended, untrained rats (Table 4). The analysis of variance detected a significant main effect of a training-induced decrease in somatomedin-C ($P<0.001$), but a significant interaction between training status and suspension treatment ($P<0.002$) showed that this decreased somatomedin-C in trained rats occurred only in the suspended animals. On the other hand, the overall mean for untrained, suspended rats was significantly greater than that for untrained, nonsuspended rats ($P<0.01$). Thus, suspension seemed to increase somatomedin-C concentration in muscles of untrained rats and decrease it in trained rats.

The mechanism underlying the training-induced increase in concentrations of somatomedin-C in plasma for the nonsuspended rats is unknown, especially since the 3 h half life of somatomedin-C is brief in relation to the 2 week interval between the end of the exercise training and the determination of somatomedin-C values. Nevertheless, it appears that suspension hypokinesia is sufficient to obliterate the positive effect of exercise training, perhaps because the suspended rats consumed less food and were undergoing a caloric and/or protein deficit. Energy and protein deficits are strongly linked to a reduction in circulating somatomedin-C (65). Unfortunately, this

Table 4

Somatomedin-C in Combined Soleus-Plantaris (Means ± SE; mU/g)

	Sedentary		Exercise	
	Saline	Dobutamine	Saline	Dobutamine
Control	270 ± 31[e]	276 ± 34[c]	271 ± 31[de]	260 ± 31 [ef]
Suspended	359 ± 31[b]	369 ± 31[a]	177 ± 31[efgh]	237 ± 29[ef]

[a](P<0.05) vs. tabled values followed by [e]
[b](P<0.05) vs. tabled values followed by [f]
[c](P<0.05) vs. tabled values followed by [g]
[d](P<0.05) vs. tabled values followed by [h]

explanation does not hold up for the untrained rats because the suspended subgroup of untrained animals also ate less than their nonsuspended counterparts, but there were no significant differences in somatomedin concentrations for these groups.

There is a correspondence in the trained groups between mean values for somatomedin-C and both mean body weights and mean muscle weights. Thus, one might speculate that the greater somatomedin values in the trained, nonsuspended rats protected them from suspension-induced loss of body and muscle mass.

The somatomedin results for the muscle extracts suggest that suspension increases somatomedin-C production or release in untrained animals but decreases it in trained animals. These data are impossible to explain based upon the limited literature available on somatomedin activity in skeletal muscle and must be considered as pilot data worthy of further experimentation.

The plantaris muscle was utilized to examine the activities of enzymes involved in the disposal of glucose. The plantaris muscle was chosen because it is characterized by a mixture of fiber types, e.g., 45% fast-twitch red, 45% fast twitch white, and 10% slow-twitch red fibers. Substrate content and the activity of enzymes involved in the disposal of glucose were examined because it was hypothesized that inactivity (H/H) would result in a lowering of the glycogen content and a decrease in the activities of enzymes responsible for glycogen synthesis, and that exercise would ameliorate these effects.

Glycogen content was determined after acid hydrolysis and subsequent fluorometric analysis of glucosyl units. This is a reliable method because the concentration of intracellular free glucose and glucose-6-phosphate are in very low concentration in resting muscle. At least three to five, twenty milligram pieces from each muscle were analyzed. The coefficient of variation for glycogen content between these pieces of the same muscle averaged 4.2%. The coefficient of variation of duplicate analysis of glucose residues in the hydrolyzate was 1.3%. Groups were pooled to test main effects since no significant interaction was found. There were no significant effects of training or drug treatment on the glycogen content. Hypodynamia/hypokinesia resulted in a significantly lower muscle glycogen content (Table 5 and Figure 18).

Table 5

Effects of the treatments on plantaris glycogen (umol/g ww) content

Exercise	16.5 (24)	Sedentary	14.6 (23)	p>0.27
Dobutamine	15.4 (23)	Saline	15.7 (24)	p> .86
Suspended	13.6 (25)	Control	17.8 (22)	p<0.02

values are means; (animals/group)

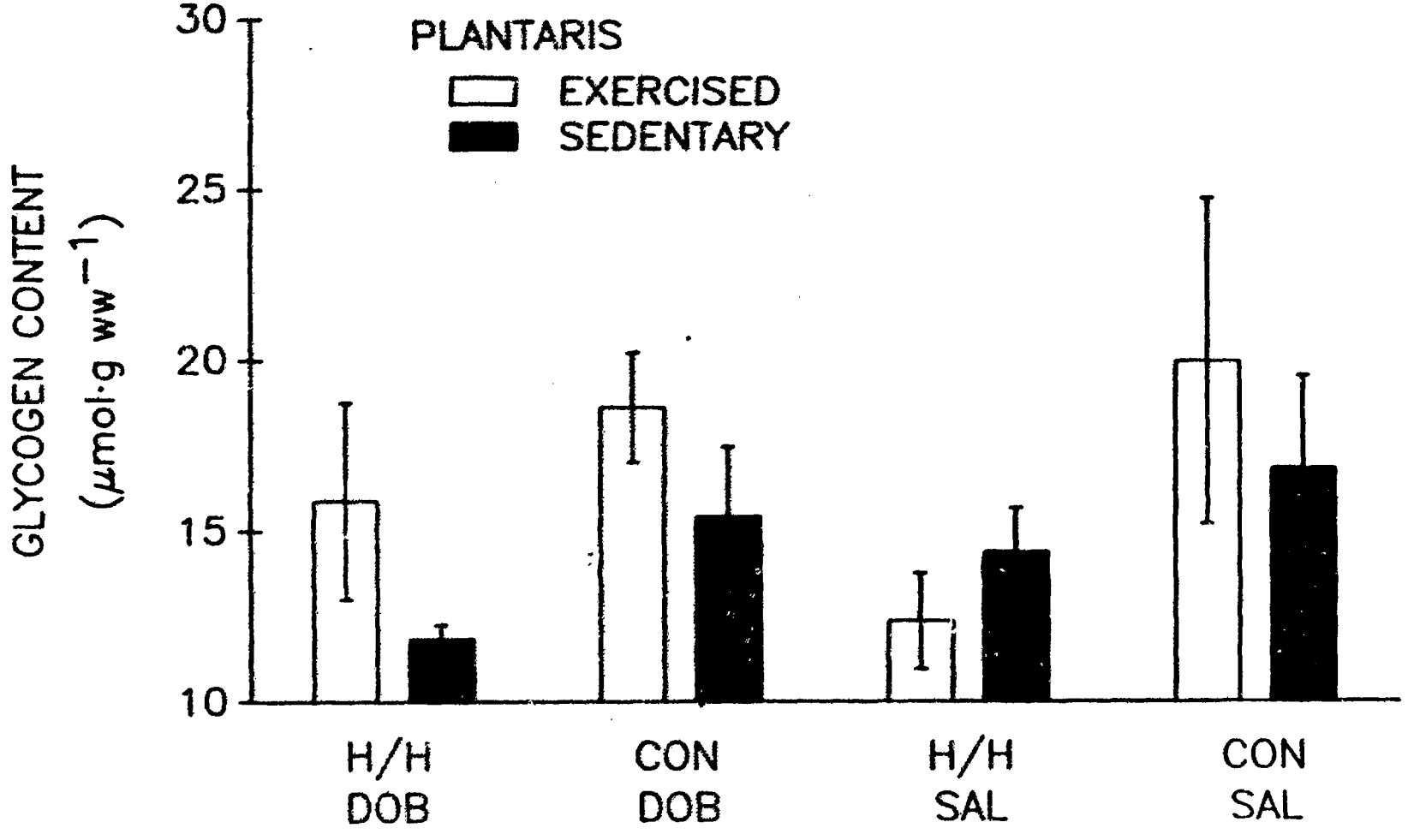

PLANTARIS
EXERCISED
SEDENTARY
GLYCOGEN CONTENT
(μmol·g ww⁻¹)
30
25
20
15
10
H/H
DOB
CON
DOB
H/H
SAL
CON
SAL

Figure 18.

Protein content of the supernatant from homogenates prepared from the plantaris muscles were analyzed using the comassie-blue, dye-binding procedure (Bradford). The activities of hexokinase and glycogen synthase were also analyzed from the supernatant allowing expression of the activity relative to protein content. The coefficient of variation for duplicate samples using this procedure averaged 2.4%. There were no significant effects of training, drug treatment or suspension on the protein content in the supernatant (Tables 6 and Figure 19).

Table 6

Effects of the treatments on plantaris homogenate supernatant protein content (mg/g ww)

Exercise	46	(24)	Sedentary	46	(23)	p>0.99
Dobutamine	46	(23)	Saline	46	(24)	p>0.97
Suspended	45	(25)	Control	48	(22)	p<0.17

values are means; (animals/group)

Hexokinase activity of the supernatant from homogenates prepared from the plantaris muscles were analyzed spectrophotometrically. The coefficient of variation for triplicate determinations of enzyme activity averaged 2.8%. There were no significant effects of training, drug treatment or suspension on the hexokinase activity in the supernatant (Table 7 and Figure 20). Similar results were found when activity was expressed relative to protein content (Table 8 and Figure 21).

Glycogen synthase activity was analyzed by following the incorporation of [14C]-uridine-diphosphate-glucose into glycogen in the absence and presence (0 and 10,000 uM) of glucose-6-phosphate, without sulfate but with 0.045 uM uridinediphosphate-glucose. The coefficient of variation of

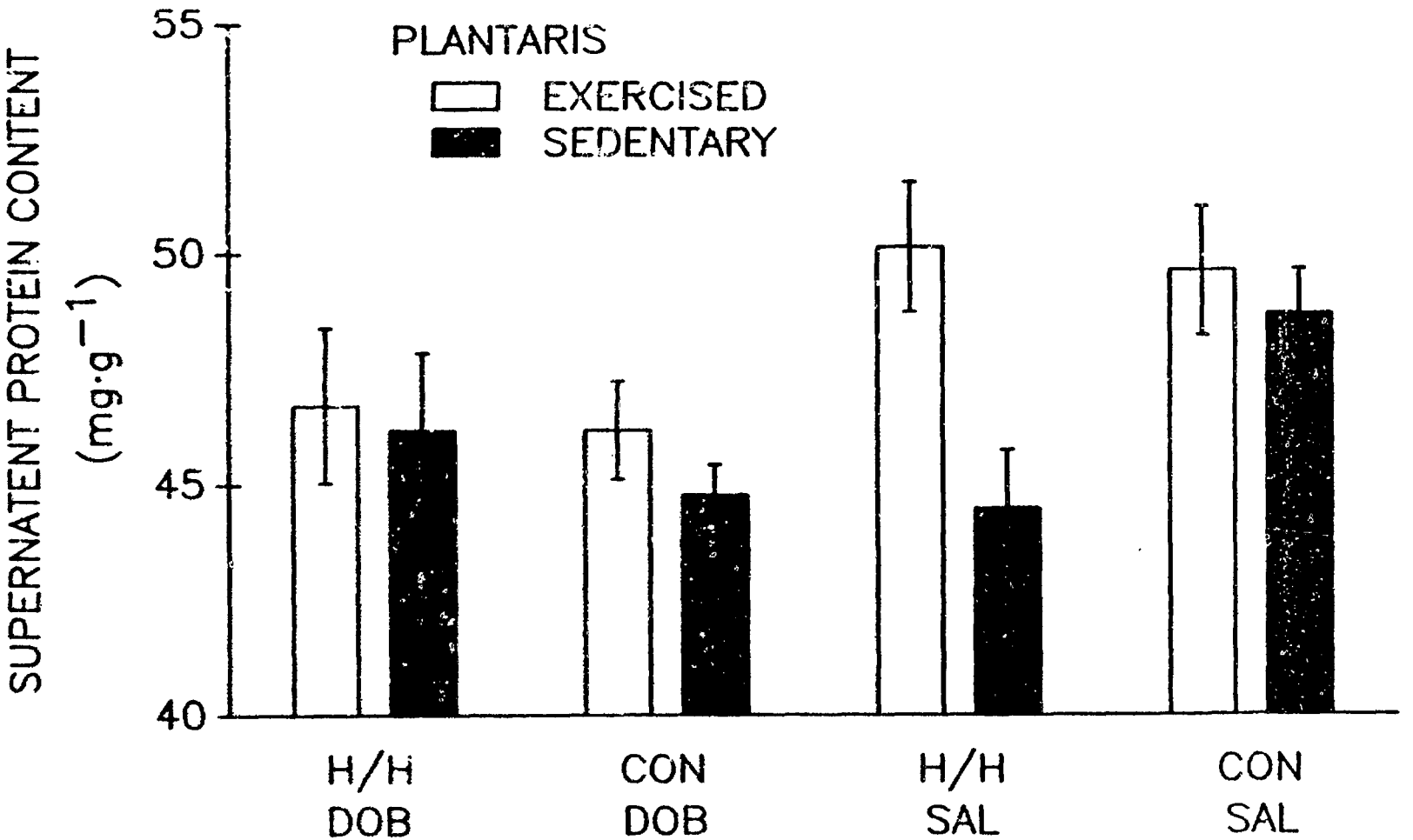

PLANTARIS
EXERCISED
SEDENTARY
SUPERNATENT PROTEIN CONTENT
(mg·g⁻¹)
55
50
45
40
H/H
DOB
CON
DOB
H/H
SAL
CON
SAL

Figure 19.

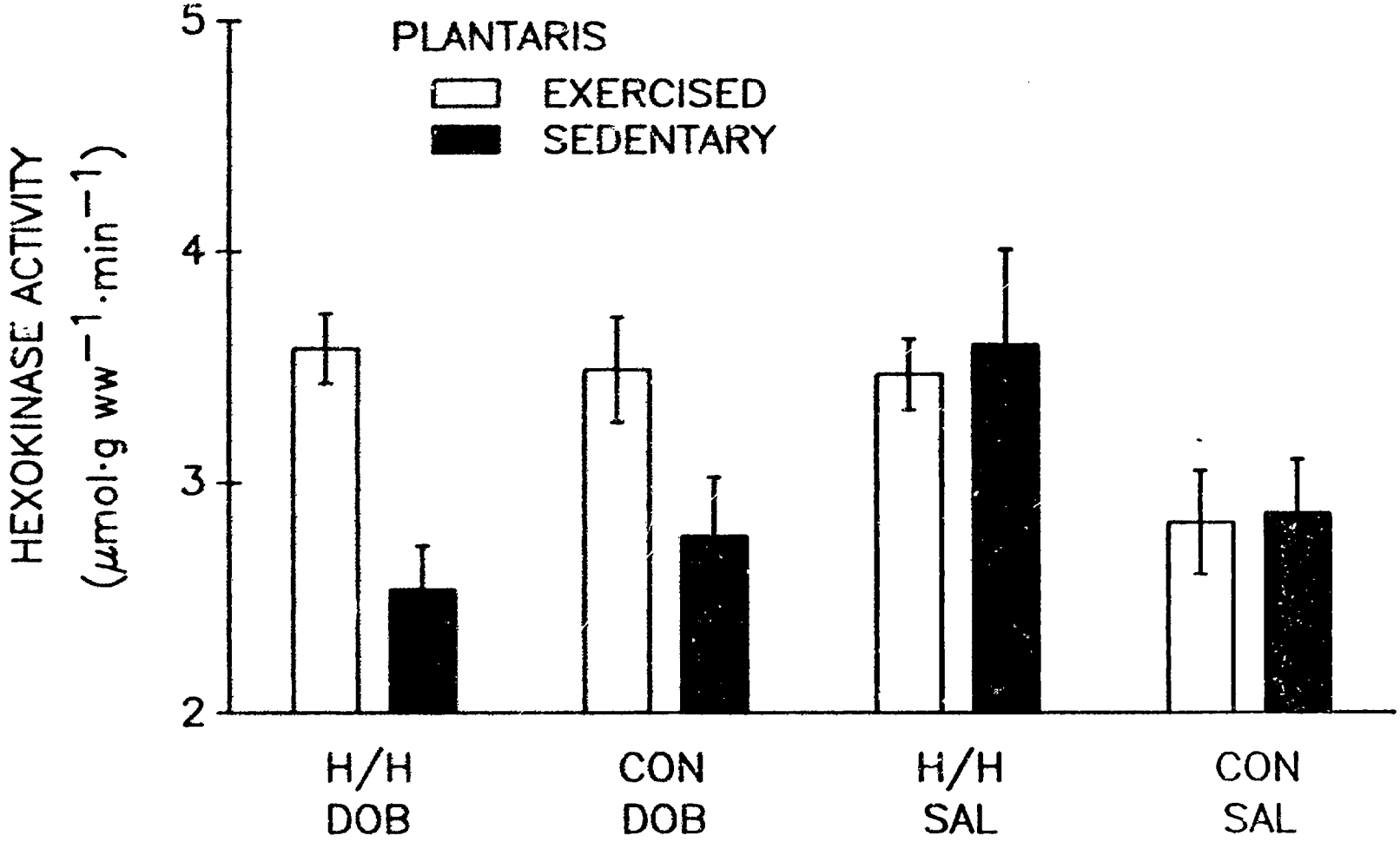

PLANTARIS
EXERCISED
SEDENTARY
HEXOKINASE ACTIVITY
(μmol·g ww⁻¹·min⁻¹)
5
4
3
2
H/H
DOB
CON
DOB
H/H
SAL
CON
SAL

Figure 20.

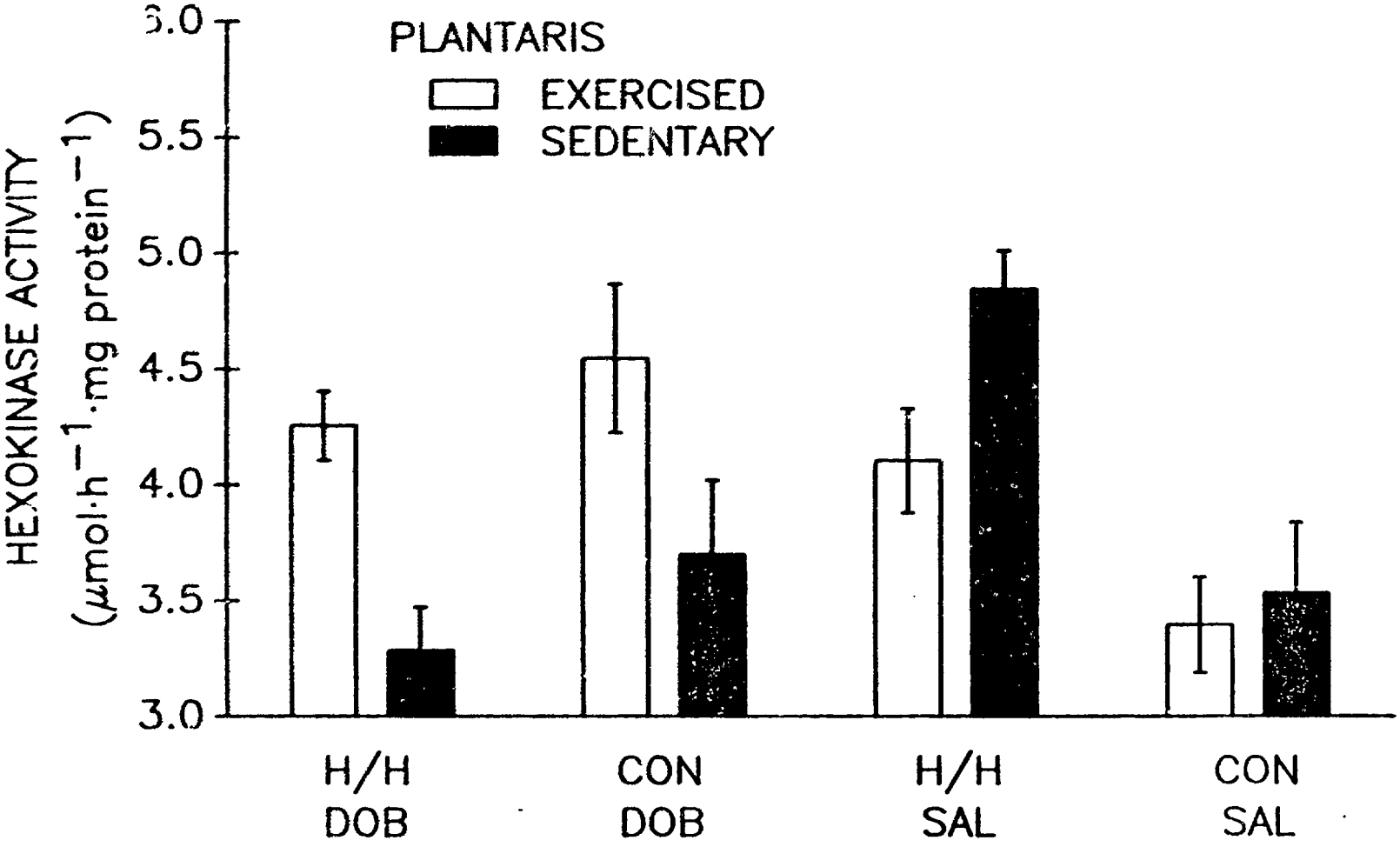

PLANTARIS
EXERCISED
SEDENTARY
HEXOKINASE ACTIVITY
(μmol·h⁻¹·mg protein⁻¹)
3.0
5.5
5.0
4.5
4.0
3.5
3.0
H/H
DOB
CON
DOB
H/H
SAL
CON
SAL

Figure 21.

Table 7

Effects of the treatments on plantaris hexokinase activity (umol/min/g ww)

Exercise	4.48 (24)	Sedentary	2.95 (23)	p>0.18
Dobutamine	4.13 (23)	Saline	4.31 (24)	p>0.29
Suspended	4.33 (25)	Control	3.05 (22)	p<0.25

values are means; (animals/group)

Table 8

Effects of the treatments on plantaris hexokinase activity (umol/mg/protein/h)

Exercise	7.14 (24)	Sedentary	3.85 (23)	p>0.28
Dobutamine	3.96 (23)	Saline	7.03 (24)	p>0.32
Suspended	6.98 (25)	Control	3.88 (22)	p<0.30

values are means; (animals/group)

duplicate analysis of supernatant from the muscles averaged less than 5%. There were no significant effects of exercise or drug treatment on glycogen synthase activity. Hypokinesia/hypodynamia significantly elevated glycogen synthase activity (Table 9 and Figure 22). When the glycogen synthase activity was expressed relative to the protein content in the supernatant, there were no significant effects of exercise or drug treatment; however, hypokinesia/hypodynamia significantly elevated glycogen synthase activity (Table 10 and Figure 23).

Table 9

Effects of the treatments on plantaris glycogen synthase activity (umol/g ww/min)

Exercise	62.3 (24)	Sedentary	59.8 (23)	p>0.95
Dobutamine	63.4 (23)	Saline	58.7 (24)	p>0.20
Suspended	65.8 (25)	Control	55.9 (22)	p<0.002

values are means; (animals/group)

Table 10

Effects of the treatments on plantaris glycogen synthase activity (nmol/mg protein/min)

Exercise	1.31 (24)	Sedentary	1.30 (23)	p>0.95
Dobutamine	1.39 (23)	Saline	1.23 (24)	p<0.07
Suspended	1.43 (24)	Control	1.18 (22)	p<0.002

values are means; (animals/group)

The glycogen synthase activity ratio was determined by dividing the enzyme activity in the absence of glucose-6-phosphate (0 uM) by the enzyme activity in the presence of the maximal concentration of glucose-6-phosphate (10,000 uM). This ratio indicates the percentage of the enzyme in the active form. There were no significant effects of exercise or drug treatment on the glycogen synthase activity ratio; however, hypokinesia/hypodynamia significantly elevated the glycogen synthase activity ratio (Table 11 and Figure 24).

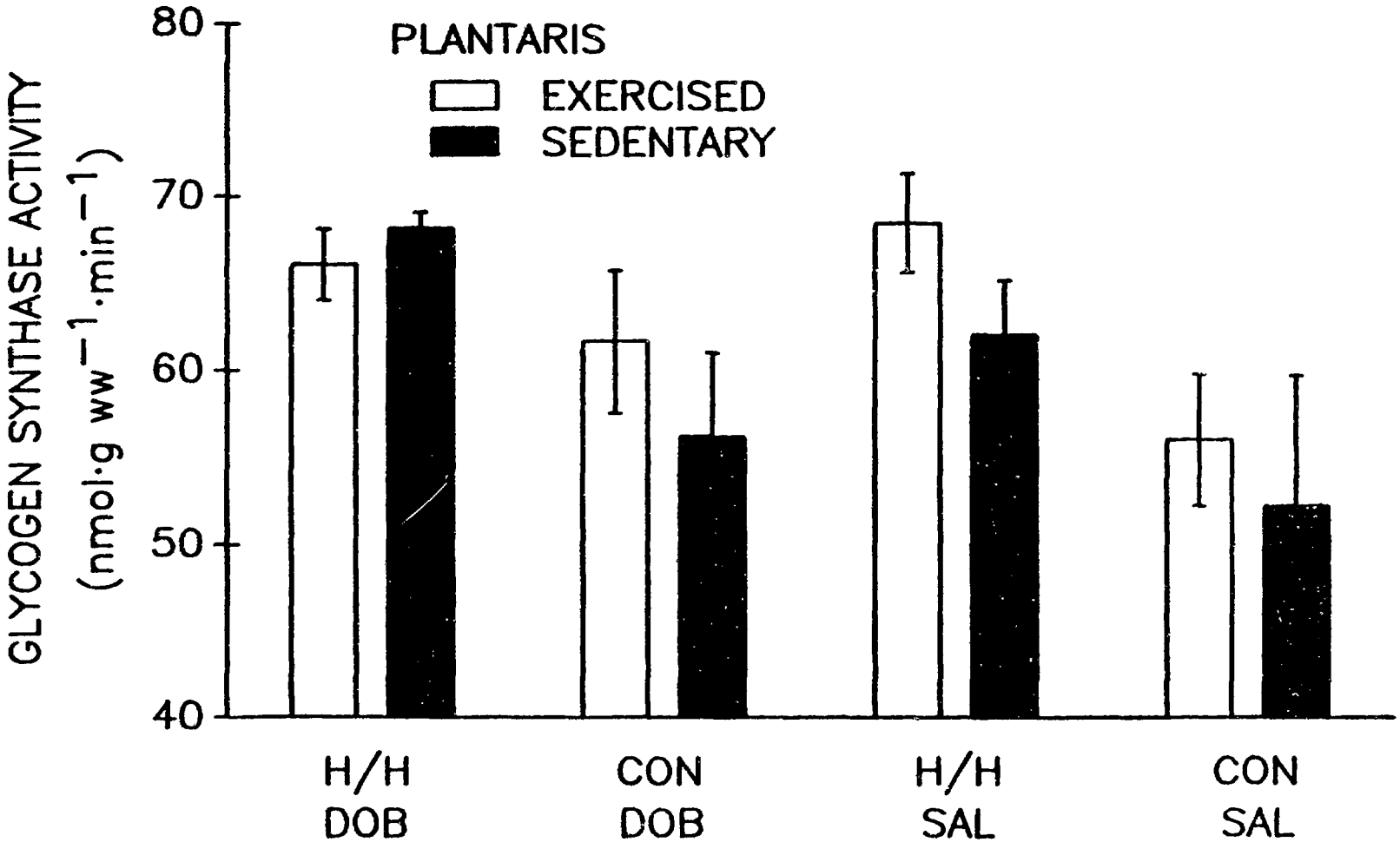

PLANTARIS
EXERCISED
SEDENTARY
GLYCOGEN SYNTHASE ACTIVITY
(nmol·g ww^{-1}·min^{-1})
80
70
60
50
40
H/H
DOB
CON
DOB
H/H
SAL
CON
SAL

Figure 22.

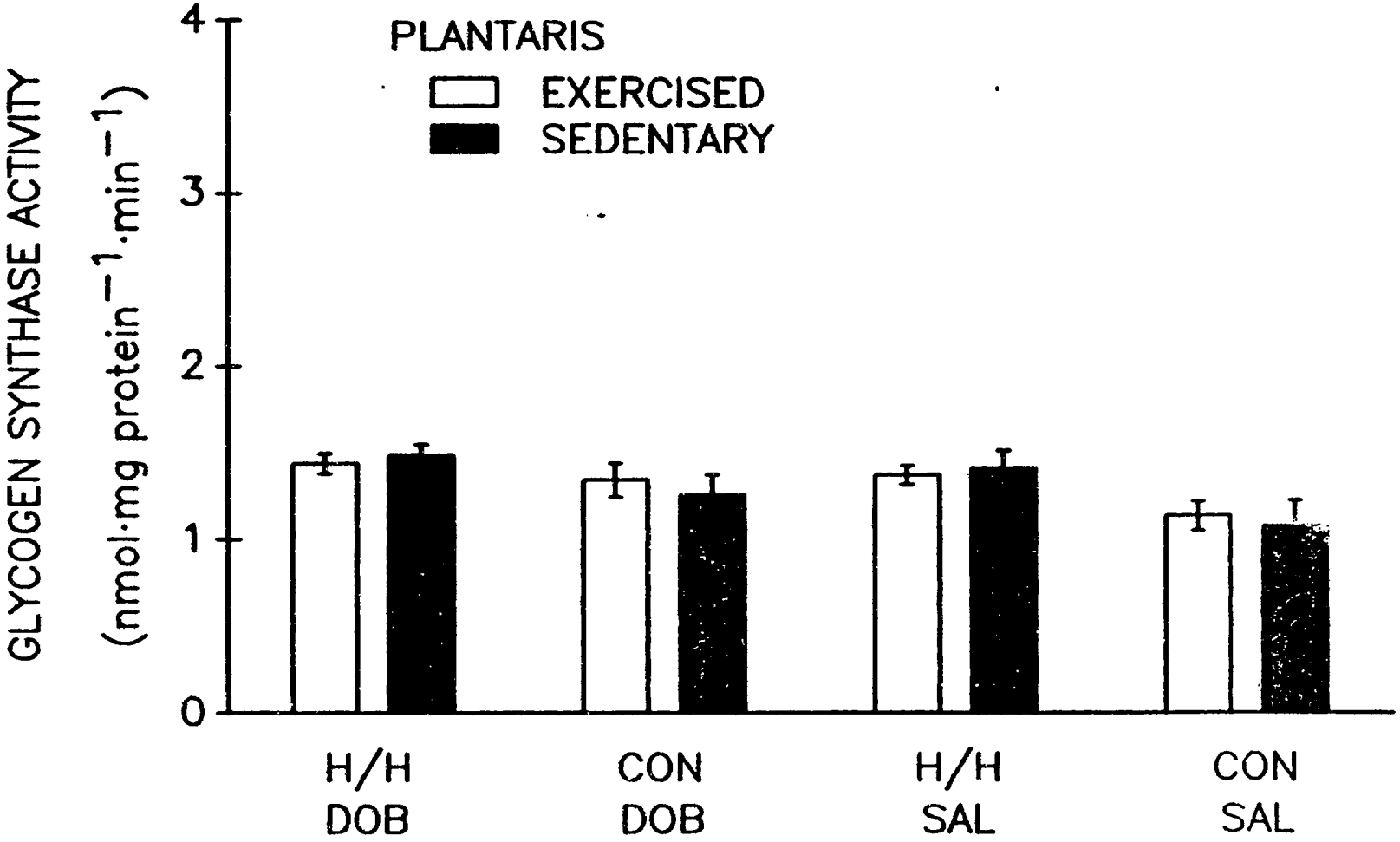

PLANTARIS
EXERCISED
SEDENTARY
GLYCOGEN SYNTHASE ACTIVITY
(nmol·mg protein^{-1}·min^{-1})
4
3
2
1
0
H/H
DOB
CON
DOB
H/H
SAL
CON
SAL

Figure 23.

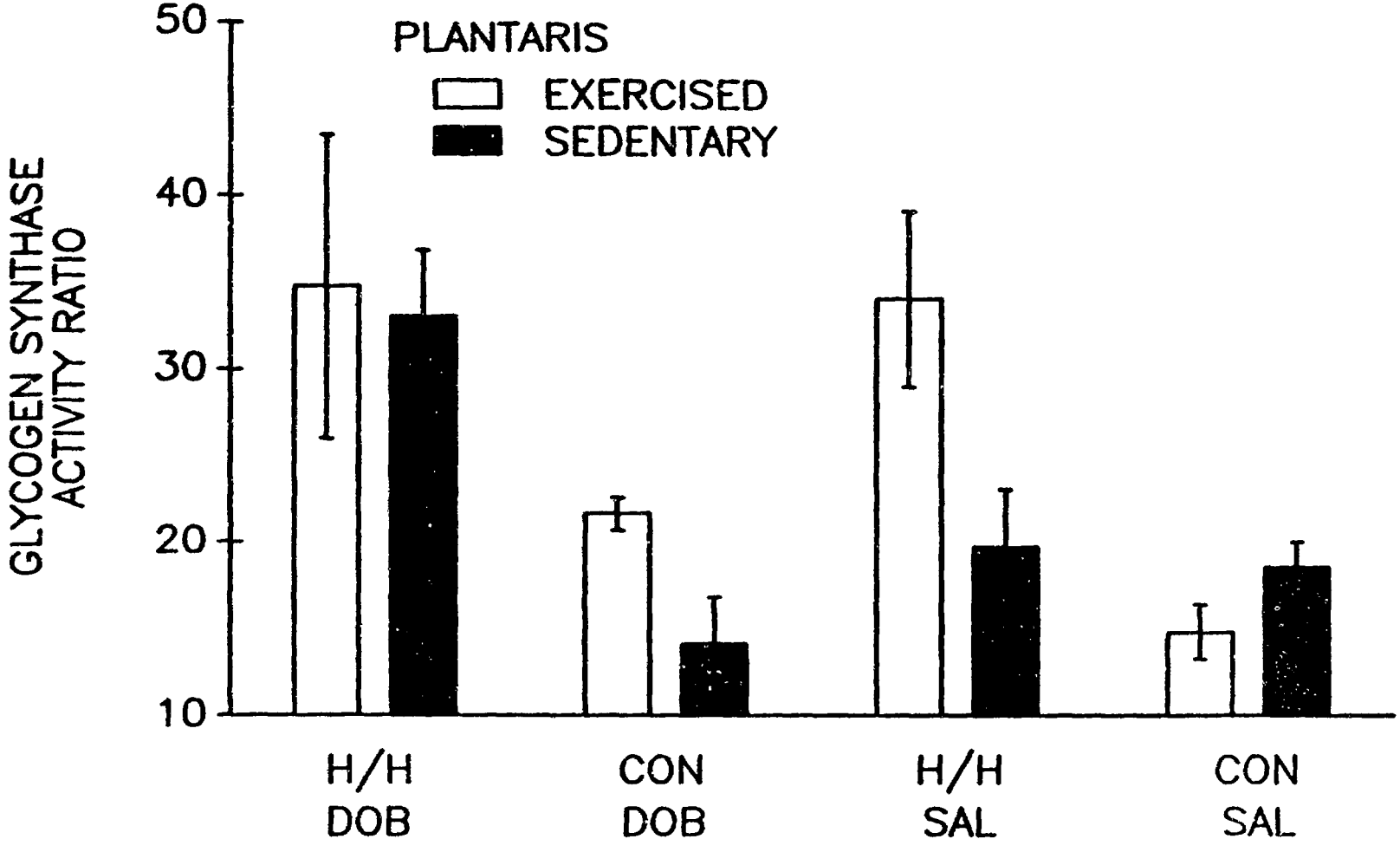

PLANTARIS
EXERCISED
SEDENTARY
GLYCOGEN SYNTHASE
ACTIVITY RATIO
50
40
30
20
10
H/H
DOB
CON
DOB
H/H
SAL
CON
SAL

Figure 24.

Table 11

Effects of the treatments on plantaris glycogen synthase activity

Exercise	26.0 (24)	Sedentary	21.2 (23)	p>0.24
Dobutamine	26.7 (23)	Saline	20.7 (24)	p<0.20
Suspended	29.8 (24)	Control	16.7 (22)	p<0.002

values are means; (animals/group)

In accordance with our hypothesis, hypokinesia/hypodynamia lowered the muscle glycogen content; however, it unexpectedly increased the activity of glycogen synthase and the activity ratio. Danforth et al., (14) observed an inverse relationship between the muscle glycogen content and the activity ratio of glycogen synthase. The present findings are in accordance with this relationship. It appears that there is a chronic adaptation in the muscle to a lowered glycogen content which also stimulates an increase in the total activity of glycogen synthase.

Exercise did not alter the effects of hypokinesia/hypodynamia on glycogen content and enzyme activity. Exercise did not increase the glycogen content of the plantaris muscle nor did it significantly affect hexokinase activity although it was 1.5-fold higher ($p<0.05$) than the sedentary group. Furthermore, exercise did not alter the effect of hypokinesia/hypodynamia on glycogen synthase activity or activity ratio. Thus, the training program used in this study was not sufficient to ameliorate the effects of hypokinesia/hypodynamia on glycogen content and glycogen synthase activity. Perhaps a more strenuous training program or a shorter period of hypokinesia/hypodynamia could produce more positive results with regard to these specific variables.

3. Cardiovascular Response and Control: These studies were carried out in Dr. Hamlin's laboratory in the College of Veterinary Medicine. Data involving heart rate and electrocardiogram response was collected by J. Hobson and Drs. Girten, Smith and Hamlin. Dr. Girten compiled and provided these results for this report. The preliminary work in baroreceptor control of heart rate and blood pressure was completed by Drs. Girten, Smith, Lash and Hamlin; J. Hobson, S. Boggs and D. Sutter provided technical assistance. Histological heart sections were prepared and analyzed by Dr. Smith and N. William and Dr. Smith provided these results for this report. Dr. Weisbrode provided support for the histological work in terms of both advice and facilities.

Electrocardiogram (ECG) components and their voltages are determined by a number of different cardiac or extracardiac factors. Cardiac determinants include the magnitude of the action potential, magnitude of the waves of depolarization and the degree of a synchronicity of simultaneous waves of depolarization. Extracardiac determinants include the volume and resistivity of blood within chambers (Brody Effect) (9) as well as volume and resistivity of pericardial sac effusates and interfaces between eccentric rings of tissues of various volumes of resistivity (Rudy-Plonsey Model) (61). One way in which these ideas could apply to the current study would be the changes that may occur in ECG waveforms as a result of changes in left ventricular volume. Increased left ventricular volume, as is often times seen with exercise training, may increase potentials and regions monitoring spread radial to the left ventricular cavity, but may decrease potentials at regions monitoring spread tangential to the cavity. These types of alterations could result in increased R wave amplitude in lead I and in some cases a decreased Q wave amplitude in lead V_{10}.

A calibration factor of 1 millvolt (mv) = 20 millimeters (mm) was used for all ECG recordings and all voltages are expressed in terms of millimeters of amplitude. Amplitude was measured from

the isoelectric line to maximal point of deflection with the exception of the QRS in lead I, which was measured peak to peak. Initial ANOVA comparisons indicated significant interaction in three of five of the ECG variables examined. When appropriate, Tukey's pairwise comparisons were used as post hoc tests. Representative waveforms are shown in Figures 25 and 26 and means ± SEM for the experimental groups are listed in Table 12.

P waves in lead I were significantly ($p \leq .05$) greater in three of four exercise groups when compared to their sedentary counterparts, with the only exception being the control dobutamine group. The P wave amplitude for the four sedentary groups and the exercised control dobutamine group were not significantly different.

Q waves in lead V_{10} were slightly greater in the four sedentary groups, but no significant ($p \leq .05$) pairwise differences were detected among any of the groups. The saline subgroups of the sedentary animals had the larger Q wave amplitude among the eight groups.

R waves in lead I were significantly ($p \leq .05$) greater in exercised suspended animals when compared to all other groups. Although all exercise groups had larger R wave amplitudes than their sedentary counterparts the only time the differences were significant were when animals were also suspended.

QRS complexes in lead AVF were significantly ($P \leq .05$) greater in three of four sedentary groups when compared to their exercise counterparts, with the only exception being the control dobutamine group. The suspended sedentary groups had significantly ($P \leq .05$) lower voltages when compared to all other groups.

S/R ratios in lead AVF were significantly ($P \leq .05$) greater in each of the four exercised groups when compared to each of the sedentary groups. The exercised saline groups were significantly

Figure 25.

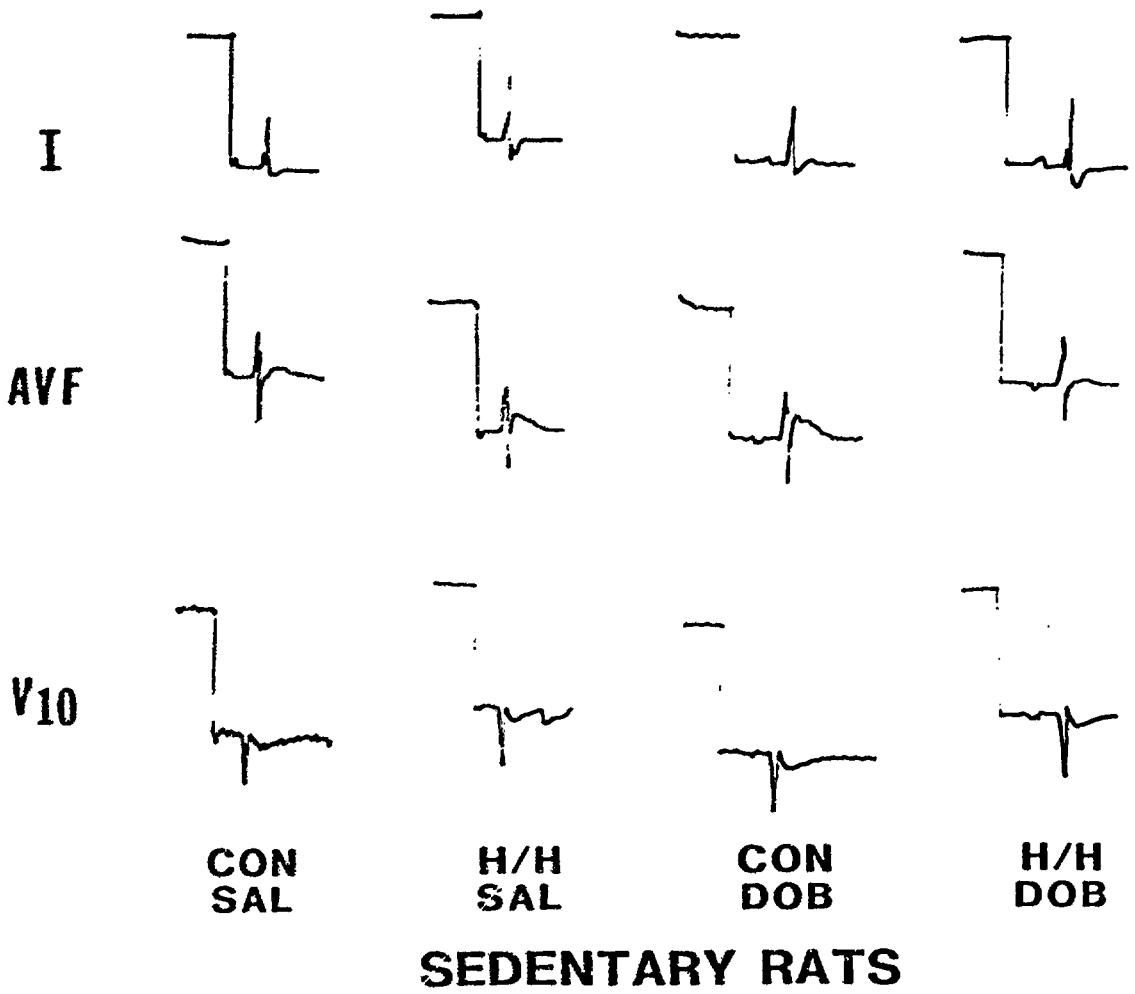


Figure 26.

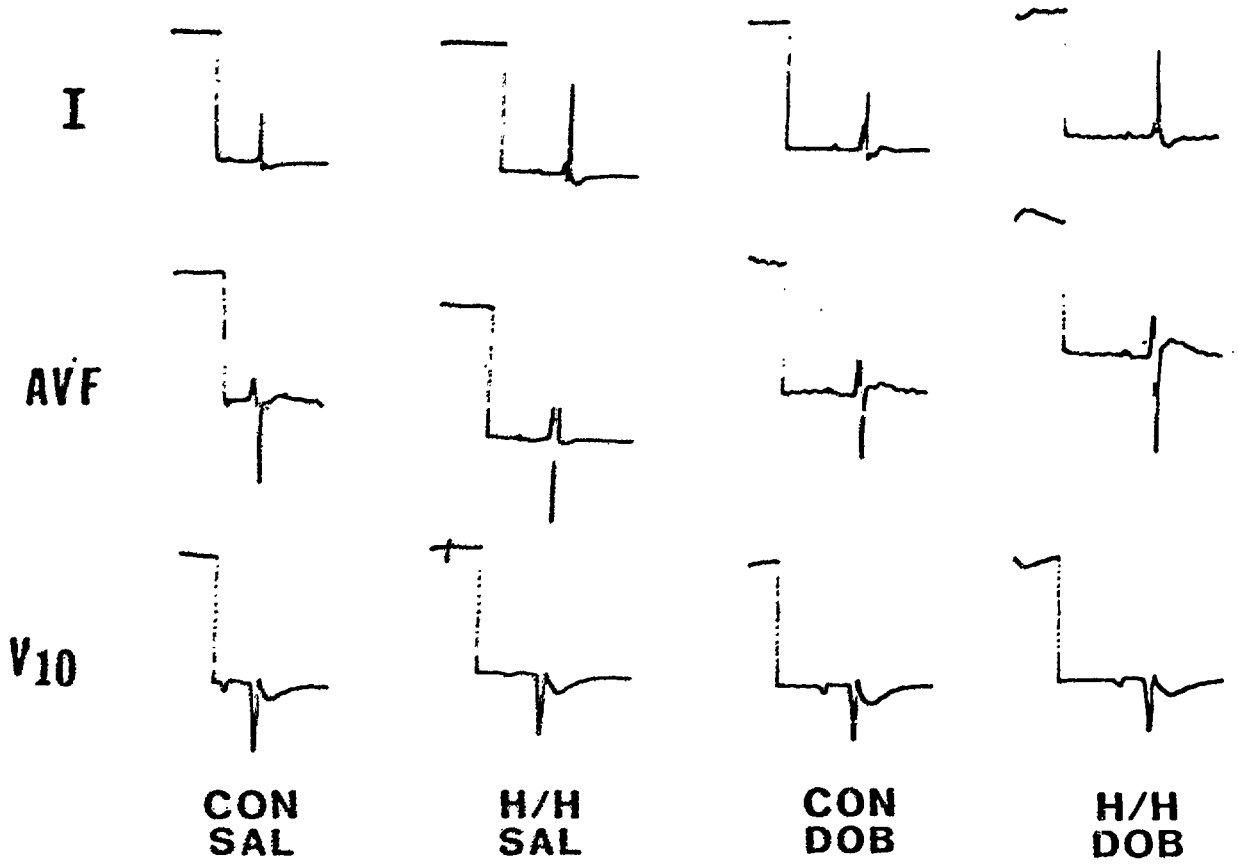


Table 12

Electrocardiographic Waveform Variables

	PI	QV_{10}	RI	QRS AVF	S/R AVF	HR
Sedentary						
Control Saline	0.9 ± 0.2	11.2 ± 1.4	8.2 ± 1.8	18.1 ± 2.1	1.0 ± 0.1	268 ± 8
Suspended Saline	0.9 ± 0.1	10.1 ± 1.6	11.4 ± 2.1	20.4 ± 2.2	0.9 ± 0.2	289 ± 7
Control Dobutamine	1.0 ± 0.2	9.6 ± 1.0	9.7 ± 1.4	16.5 ± 1.8	0.8 ± 0.1	259 ± 8
Suspended Dobutamine	1.2 ± 0.1	9.4 ± 1.0	12.0 ± 1.7	22.7 ± 2.3	0.7 ± 0.1	274 ± 10
Exercise						
Control Saline	1.6 ± 0.2	9.5 ± 1.4	10.2 ± 1.3	14.2 ± 1.3	4.4 ± 0.1	219 ± 6
Suspended Saline	1.9 ± 0.3	8.7 ± 0.9	17.1 ± 2.2	13.6 ± 1.8	2.5 ± 0.2	243 ± 9
Control Dobutamine	1.2 ± 0.1	9.3 ± 1.4	10.8 ± 1.4	15.0 ± 1.7	2.2 ± 0.1	208 ± 7
Suspended Dobutamine	2.2 ± 0.2	8.8 ± 1.3	16.8 ± 1.3	13.4 ± 1.4	2.4 ± 0.1	236 ± 9

($P \leq .05$) greater than the exercised dobutamine groups, but there were no significant differences among the sedentary groups.

Heart rate was significantly ($P \leq .05$) higher in each of the sedentary groups when compared to each of the exercised groups. The exercised suspended groups were significantly ($P \leq .05$) greater than the exercised control groups and although the sedentary suspended groups were higher than the sedentary control groups these differences were not significant.

Heart rates for each of the four sedentary groups were significantly ($P \leq .05$) higher than each of the exercised groups. The exercised suspended animal had significantly ($P \leq .05$) greater heart rates than the exercised control animals and although this same trend was observed in the sedentary groups the differences were not significant.

Exercised trained rats had larger P and R waves in lead I than did sedentary animals and this agrees with previous results for mice that had been swim trained (38). The differences between exercise and sedentary groups for QRS complex and S/R ratio apparent in the present study were not observed in the swim trained mice. Some of these changes may have resulted from lower body weights of the exercised animals or from changes in resistivity of blood resulting from changed hematocrit in these groups (9). There is no information available on ECG changes in rodents after suspension or dobutamine treatments. The observed change in the QRS complex as a result of suspension may have also been related to weight change or, as is suggested by the Rudy-Plonsey model, it could be due to alterations in percentages of fat and/or muscle (61). Lower heart rates in the exercised groups is typical of the bradycardia observed in many endurance trained animals (38). Although there is not information available with regard to increased heart rate in suspended animals it is interesting to note that the change in myosin isoenzyme patterns towards a pattern similar to that seen in

hypertensive rats (64) were also observed in the suspended groups.

Baroreflex control is determined by measurement of arterial baroreflex sensitivities or the Δ heart rate/ Δ mean arterial pressure. Preliminary work on development of methods that would allow for this type of measurement during exercise, dobutamine and suspension treatments were conducted. Response to infusions of drugs such as phenylephrine and nitroprusside after, or during, the above treatments was also of interest.

Initial surgical preparations involved simultaneous cannulation of the carotid artery and jugular vein with a straight 23-guage polyethylene cannulae. Although animals survived the surgery well and regained full mobility many of the cannulae become partially or completely obstructed. Since the carotid cannulae were above the level of the renal arteries attempts to flush the catheter could result in infarcted kidneys. Some records of heart rate and blood pressure changes with drug infusion at rest and during exercise were obtained during this phase of the investigation, but since the problems with the carotid cannulation procedure persisted, an alternate surgical technique was explored. Samples of the records obtained with this initial surgical preparation, with phenylephrine at rest and during exercise are shown in Figures 27 and 28.

The second surgical preparation involved simultaneous cannulation of the abdominal aorta and the jugular vein. Dr. J. Weeks of the Cardiovascular Disease Research Group of The Upjohn Company has developed detailed cannula techniques and instrumentation for this type of chronic implantation for rats. D. Sutter, a research biologist from Upjohn, provided technical advice and assistance during this phase. Since Upjohn had used this technique for chronic implants up to several months we were interested in determining if exercised trained animals could undergo this procedure and continue training afterward without

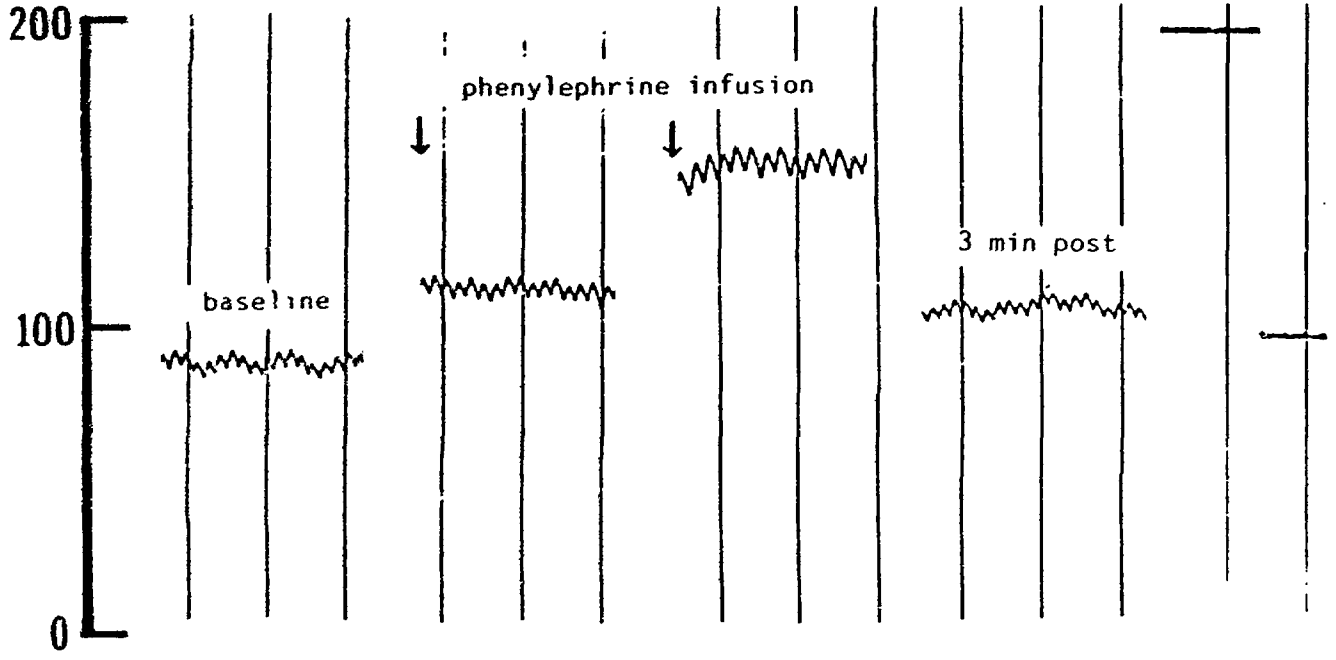


Figure 27.

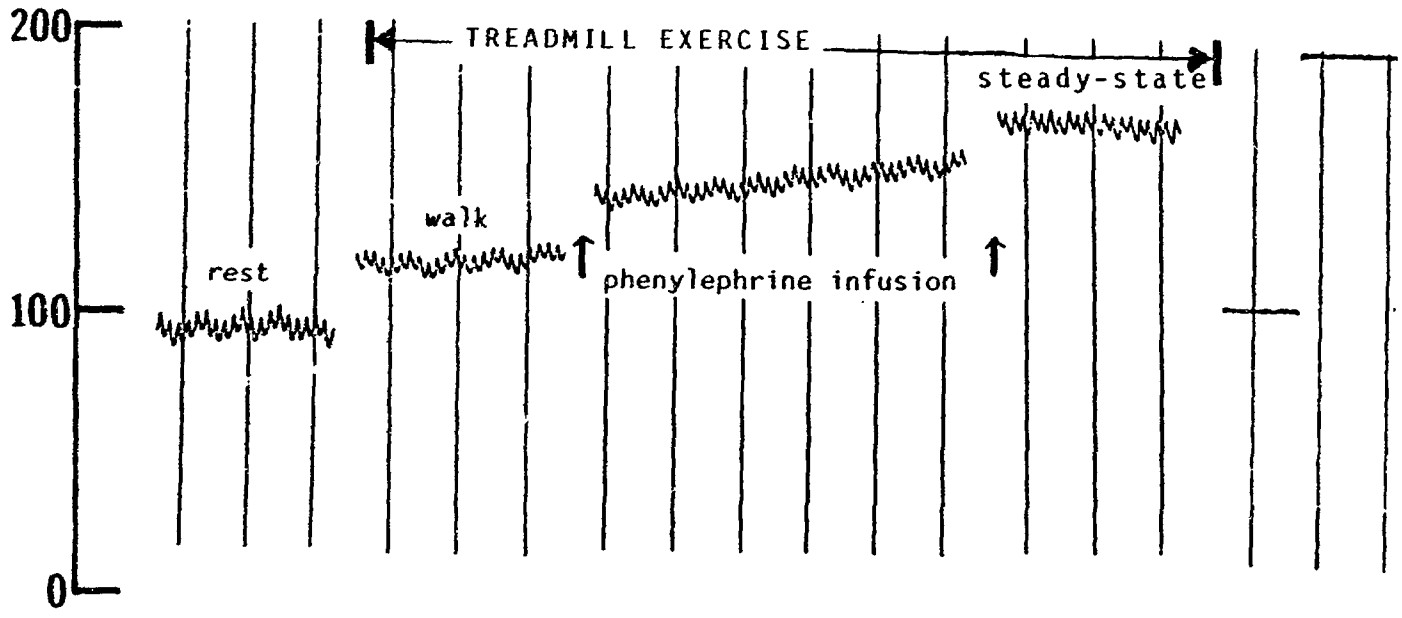

200
100
0
TREADMILL EXERCISE
steady-state
walk
rest
phenylephrine infusion

Figure 28.

significant detraining taking place. Details involving these cannulation techniques are included in Appendix D.

Animals that had been exercise trained for approximately fourteen weeks were able to start on a modified exercise program two days after the abdominal aorta and jugular cannulas had been surgically implanted. These animals were exercised for an additional four weeks and soleus citrate synthase levels and resting heart rates indicated that most of the gains seen after fourteen weeks of training remained after surgery and the four week program of modified exercise. Citrate synthase levels of the cannulated rats were approximately 84% of previously trained animals and still significantly ($P \leq .05$) greater than sedentary animals, but averaged approximately 10 bpm faster than trained animals who had not undergone surgery. Although small sample sizes (n=5 per group) were used, these comparisons provide valuable information for future investigations utilizing these techniques and treatments.

Histological heart sections indicated that none of the treatments produced any pathological problems that would be identified with Massons trichrome stain. Views from a heart that was representative of all groups are shown in Figures 29. A heart section from this study is compared to a comparable section of a rat heart that was infarcted by coronary artery ligation in Figures 30 and 31. Large areas of fibrotic tissue stained light blue as a result of the infarction compared to the very small amounts of light blue stained collagen that surrounds major vessels in the healthy heart.

Cross-sectional ventricular wall area was computed by first tracing the epicardial surface of the left ventricle in a clockwise manner, then tracing the endocardial surface of the left ventrical in counterclockwise fashion. When the endocardial surface is traced in the opposite direction it is subtracted from the area generated by tracing epicardial circumference to yield

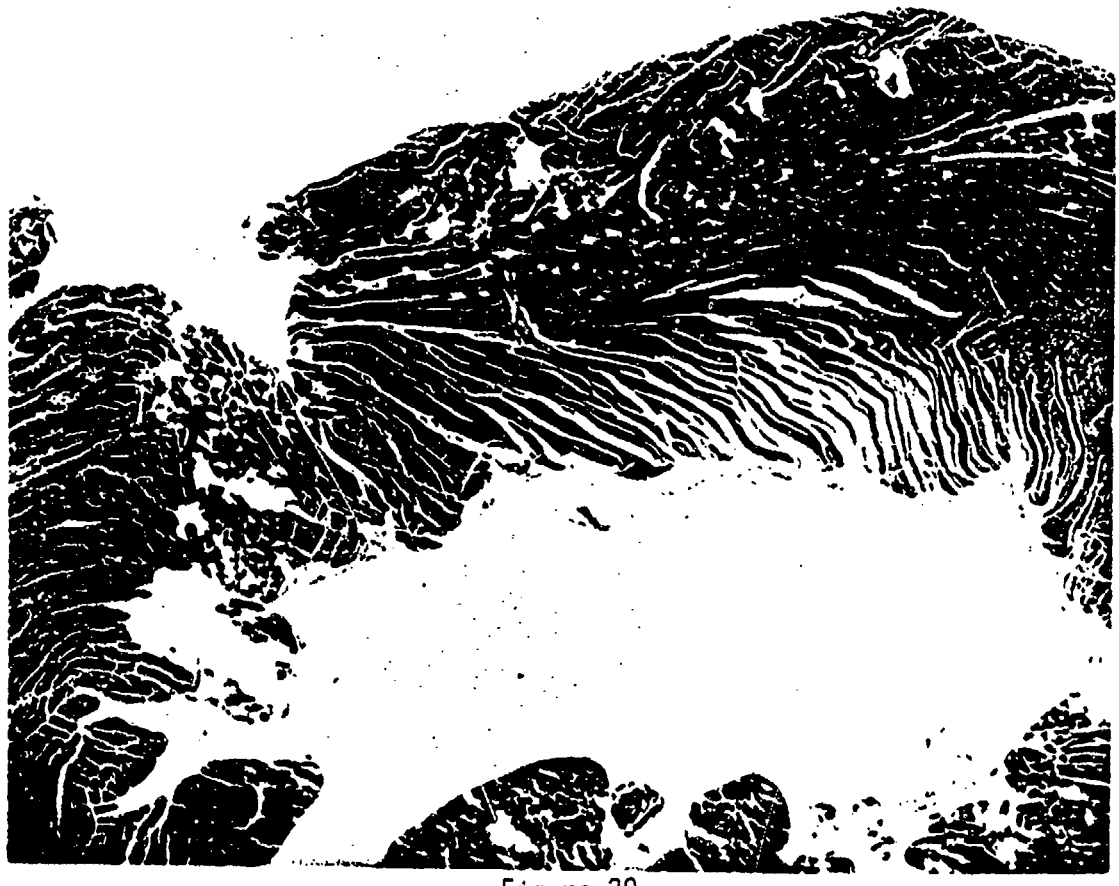


Figure 29.



Figure 30.

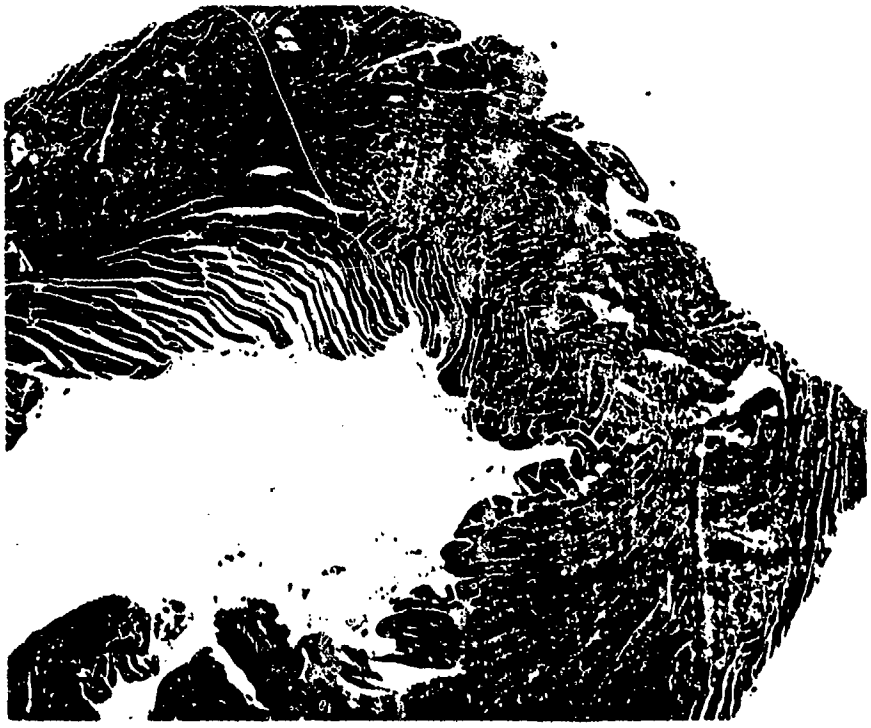


Figure 31.



the true area of the left ventricle minus the cavity. Coefficient of variation for tracing was 0.9%.

Group means (± SEM) for left ventricular wall area, body weight, heart weight/body weight ratio and area/body weight ratio are given in Table 13. Initial ANOVA procedures on the area and two ratios indicated that there was no significant ($P \leq .05$) interaction of suspension with the other treatments for these three variables so groups were pooled with regard to this treatment. Post hoc analysis showed that there were no significant ($P \leq .05$) pooled group differences for wall area or heart weight/body weight ratio, but there was a significant effect for the area/body weight ratio with regard to dobutamine treatment. The pooled dobutamine control groups had a significantly ($P \leq .05$) lower area/body weight ratio when compared to the other groups. This indicates that when the lower body weights of the exercised animals were taken into account a difference that was not otherwise apparent became significant.

E. Conclusion and Recommendations

Data obtained in this study provides valuable information with regard to hypokinesia/hypodynamia suspension, exercise training and dobutamine treatment and their interactive effects on each other. Some of the results from this investigation into the areas of biochemical adaptations, metabolic control factors and cardiovascular response and control serve to support established information involving the treatments; however, many of the results represent findings that had not been previously investigated. Although much of the data obtained supported our original hypotheses, there were a number of instances where unexpected significant differences were found or where the results failed to support initial expectations. Overall, the information provided as a result of this investigation gives insight into mechanisms responsible for many of the descriptive changes described by previous studies with regard to the treatments examined.

Table 13

Left Ventricular Wall Area and Related Variables

	LV WALL Area	HRT WT	BODY WT	HW/BW RATIO	AREA/BW Ratio
Sedentary	(mm^2)	(grams)	(grams)	(x100)	(mm^2/g)
Saline Control	56.2 ± 3.0	1.371 ± 0.053	436 ± 9	3.14 ± 0.09	0.130 ± 0.009
Saline Suspended	57.3 ± 2.6	1.346 ± 0.033	381 ± 6	3.53 ± 0.09	0.151 ± 0.007
Dobutamine Control	48.8 ± 4.5	1.350 ± 0.034	417 ± 9	3.24 ± 0.09	0.116 ± 0.009
Dobutamine Suspended	53.8 ± 4.6	1.425 ± 0.040	397 ± 7	3.60 ± 0.13	0.135 ± 0.010
Exercised					
Saline Control	50.7 ± 1.8	1.326 ± 0.047	377 ± 8	3.52 ± 0.09	0.135 ± 0.004
Saline Suspended	49.6 ± 2.9	1.144 ± 0.027	346 ± 7	3.31 ± 0.03	0.143 ± 0.008
Dobutamine Control	38.1 ± 3.2	1.221 ± 0.062	364 ± 6	3.35 ± 0.13	0.100 ± 0.006
Dobutamine Suspended	47.9 ± 3.2	1.195 ± 0.019	344 ± 7	3.48 ± 0.10	0.143 ± 0.011

Most of the results obtained involving the biochemical adaptations of oxidative and antioxidant training enzymes supported our original hypotheses. Exercise training did increase these enzymes and suspension did decrease them. Although suspension often resulted in more severe loss of activity in exercised trained animals, the final enzyme activity levels for the exercised groups was as high or higher than the presuspension levels of untrained animals. The short-term (14 days) dobutamine treatment did not often change the enzyme status in control animals, but it did appear to have positive effects in terms of helping maintain enzyme activity with suspension. Myosin isoenzyme changes from the exercise treatment were expected, but the changes observed from suspension had not been previously investigated and had not been anticipated.

Previous information on somatomedin-C and the treatments was limited. Results obtained for plasma somatomedin-C indicate that suspension hypokinesia was sufficient to obliterate the positive effect of exercise training. Significant, but opposite effects in the muscle somatomedin-C levels had not been anticipated. Muscle glycogen content was lowered as a result of the suspension treatment as expected; however, it unexpectedly increased the activity of glycogen synthase and the activity ratio. These findings support the previously observed inverse relationship between muscle glycogen content and the activity ratio of glycogen synthase.

The larger P and R waves and resting bradycardia observed in the exercised trained animals had been expected, but the differences in the QRS complex and the S/R ratio had not been anticipated. Suspension induced changes in the QRS complex were not surprising in light of the changes in body weight and the increased heart rate observed with this treatment may be related to observed changes in the myosin isoenzymes. Preliminary information on measurement of baroreflex control during these types of treatments should prove valuable for future

investigations. Examination of the stained histological heart sections indicated that all groups showed similar characteristics and that there were no obvious treatment related problems. When lower body weights of exercised animals was taken into account in the ventricular area/body ratio, lower values for the dobutamine treated groups became apparent. This was the only significant difference identified for the ventricular area related variables.

Many of the results obtained during this investigation suggest that exercise training prior to suspension hypokinesia/ hypodynamia can be helpful. Additionally, even though the 14 day dobutamine treatment did not induce tremendous changes in and of itself, it did appear beneficial in maintaining some of the positive effects induced by exercise training prior to the suspension. The most notable exception to this was seen in the plasma somatomedin-C.

The findings presented here provide important information that may be useful in help reduce some deconditioning changes commonly seen with exposure to simulated or actual weightlessness. Continued investigation in this area will clarify some of the questions and unexpected findings that resulted from this study. Specific follow-on programs for dobutamine might involve examination of different modes of short and long term administration of dobutamine in rats and delineation of the drug's usefulness in ground-based hypokinesia studies with nonhuman primates. Further exercise studies could examine effects of various forms of increased activity during simulated weightless conditions.

F. References

1. Banister, E.W., Griffiths, J. Blood Levels of Adrenergic Amines During Exercise, Journal of Applied Physiology, Vol. 33, pp. 674-676, 1972.

2. Bartels, R.L., D.R. Lamb, J. Snook, V. Vivian, K. Rinehardt, and K. Ward. Anabolic Adaptation to Heavy Resistance Training While Consuming a High Energy Diet. Exercise Physiology Laboratory, The Ohio State University: manuscript in preparation, 1988.

3. Bedford, T.G., Tipton, C.M., Wilson, N.C., Opplinger, R.A., Gisolfi, C.V. Maximum Oxygen Consumption of Rats and its Changes with Various Experimental Procedures. Journal Applied Physiology: Respiratory, Environmental and Exercise Physiology, Vol. 46, no. 6, pp. 1278-1283.

4. Bello, M.A., Roy, R.R., Edgerton, V.R. Hindlimb Suspension Effects on the Morphologic and Metabolic Properties of Rat Medical Gastrocnemius. The Physiologist, Vol, 28, no. 4, p.315, 1985.

5. Blethen, S.L., W.H. Daughaday, V.V. Weldon. Kinetics of the Somatomedin-C/Insulin-Like Growth Factor I: response to exogenous growth hormone (GH) in GH-deficient children. Journal Clinical Endocrine Metabolism, Vol. 54, pp. 986-1000, 1982.

6. Blomquist, C.G. Cardiovascular Adjustments to Gravitational Stress. In: Handbook of Physiology, Section 2 The Cardiovascular System. Vol. 3, Bethesda MD: American Physiological Society, pp. 1025-1063, 1983.

7. Booth, F.W. and Golnick, P.D. Effects of Disuse on the Structure and Function of Skeletal Muscle, Medicine and Science in Sports and Exercise, Vol. 15, no. 5, pp. 415-420, 1983.

8. Brodowicz, G.R., Girten, B.E., Lamb, D.R. Exercise Training Attenuates Myocardial Creatine Kinase Release After Isoproterenol Administration in the Rat. Medicine and Science in Sports and Exercise, Vol. 18, no. 2, pp. 594, 1986.

9. Brody, D.A. A Theoretical Analysis of Intracavity Blood Mass Influence on Heart Lead Relationship. Circulation Research, Vol. 4, p. 721, 1956.

10. Chiasson, J.L., Shikama, H., Chu, D.T.W., Exton, J. H. Inhibitory Effect of Epinephrine on Insulin-Stimulated Glucose Uptake by Rat Skeletal Muscle. Journal Clinical Investigation, Vol. 68, pp. 706-713, 1981.

11. Cohen, G., Dembier, D., Marcus, J. Measurement of Catalase Activity in Tissue Extracts. Analytical Biochemistry, Vol. 34, p. 30, 1970.

12. Cohen, K.L., and S.P. Nissley. The Serum Half-Life of Somatomedin Activity: Evidence for Growth Hormone Dependence. Acta Endocr., Vol. 83, pp. 243-258, 1976.

13. Costill, D.L., Fink, W.J., Van Handel, P.J. Analytical Methods for the Measurement of Human Performance, pp. 109-111, Ball State University, 1979.

14. Danforth, W. Activation glycogen synthase and skeletal muscle. Journal of Biological Chemistry, Vol. 240, p. 1138, 1963.

15. Davidson, W.R., Banerjee, S.P., Liang, C.S. Dobutamine-Induced Cardiac Adaptations: Comparisons with Exercise-Trained and Sedentary Rats. American Journal Physiology, Vol. 250, pp. H725-H730, 1986.

16. Dietlein, L.F. Skylab: A Beginning. In: Biomedical Results from Skylab, R.S. Johnson and L.F. Dietlein (Eds.) Washington DC: NASA, pp. 408-418, 1977.

17. Dohm, G.L., Beecher, G.R., Stephenson, T.P., Womack, M. Adaptations to Endurance Training at Three Intensities of Exercise. Journal Applied Physiology: Respiratory, Environmental and Exercise, Vol. 42, no. 5, pp. 753-757, 1977.

18. Fishbein, M.C., Maclean, D., Maroko, P.R. Experimental Myocardial Infarction in the Rat. American Journal of Pathology, Vol. 90, pp. 57-70, 1978.

19. Flynn, D.E. and Max, S.R. Effects of Suspension Hypokinesia/Hypodynamia on Rat Skeletal Muscle. Aviation, Space and Environmental Medicine, pp. 1065-1069, 1985.

20. Furlanetto, R.W., L. Underwood, J.J. Van Wyuk, and A.J. D'Ercole. Estimation of Somatomedin-C Levels in Normals and Patients with Pituitary Disease by Radioimmunoassay. J. Clin. Invest., Vol. 60 pp. 648-657, 1977.

21. Hamlin, R.L., Klepinger, W.L., Gilpin, K.W., Smith, C.R. Autonomic Control of Heart Rate in the Horse, American Journal of Physiology, Vol. 222, no. 4, pp. 976-978, 1972.

22. Hauschka, E., Roy, R.R., Ederton, V.A. Hindlimb Suspension Effects on the Morphologic and Metabolic Properties of the Rat Soleus. The Physiologist, Vol. 28, no. 4, p. 315, 1985.

23. Heath, G.W., Gavin, III, J.R., Hinderliter, J.M., Haybery, J.M., Bloomfield, S.A., Holloszy, J.O. Effects of Exercise and Lack of Exercise on Glucose Tolerance and Insulin Sensitivity. Journal of Applied Physiology, Vol. 55, pp. 512-517, 1983.

24. Herbison, G.J. and Talbot, J.M. Muscle Atrophy During Space Flight: Research Needs and Opportunities. The Physiologist, Vol. 28, no. 6, pp. 520-527, 1985.

25. Hoh, J.F.Y. Myosin Isoenzymes in Skeletal and Cardiac Muscles. In: Cross Bridge Mechanisms in Muscle Contraction, Ed.: H. Sugi and G.H. Pollack, University Park Press, Baltimore, pp. 488-498, 1979.

26. Holder, A.T., E.M. Spencer and M.A. Preece. Effect of bovine growth hormone and a partially pure preparation of somatomedin on various growth parameters in hypopituitary dwarf mice. J. Endocr., Vol. 89, pp. 275-282, 1981.

27. Holloszy, J.O. Biochemical Adaptations in Muscle: Effects of Exercise on Mitochondrial Oxygen Uptake and Respiratory Enzyme Activity in Skeletal Muscle. The Journal of Biological Chemistry, Vol. 242, no. 9, pp. 2278-2282, 1967.

28. Horton, E.S. Role of Environmental Factors in the Development of Non-insulin Dependent Diabetes Mellitus. American Journal of Medicine, Vol. 72, pp. 32-40, 1983.

29. Hung, J., Goldwater, D., Convertino, V.A., et al. Mechanisms for Decreased Exercise Capacity After Bed Rest in Normal Middle-Aged Men. The American Journal of Cardiology, Vol. 51, pp. 344-348, 1983.

30. Ivy, J.L., Holloszy, J.O. Persistent Increase in Glucose Uptake by Rat Skeletal Muscle Following Exercise, American Journal of Physiology, Vol. 241, pp. C200-203, 1981.

31. Ivy, J.L., Young, J.C., McLane, J.A., Fell, R.D., Holloszy, J.O. Exercise Training and Glucose Uptake by Skeletal Muscle. Journal of Applied Physiology, Vol. 55, pp. 1393-1396, 1983.

32. Ivy, J.L., Sherman, W.M., Cutter, C.L., Kutz, A.L., Van Dyke, J. Effect of Exercise Training and Diet on Insulin Stimulated Glucose Uptake and Disposal in the Obese Zucker Rat, American Journal of Physiology, in press, 1986.

33. Jaspers, S.R. and Tischer, M.E. Atrophy and Growth Failure of Rat Hindlimb Muscle in Tail-Cast Suspension. Journal of Applied Physiology, Vol. 52, pp. 1475-1479, 1984.

34. Jenkins, R.R. The Role of Superoxide Dismutase and Catalase in Muscle Fatigue, In Biochemistry of Exercise, International Series on Sports Sciences, Vol. 13, pp. 467-471, 1983.

35. Jenkins, R.R., Friedland, R., Howard, H.H. The Relationship of Oxygen Uptake to Superoxide Dismutase and Catalase Activity in Human Skeletal Muscle. International Journal of Sports Medicine, Vol. 5, pp. 11-14, 1984.

36. Kahn, C.R. Insulin Resistance, Insulin Insensitivity, and Insulin Unresponsiveness: A Necessary Distinction. Metabolism, Vol. 27 (Suppl. 2), pp. 1893-1902, 1978.

37. Kanter, M.M. The Efforts of Chronic Aerobic Exercise Training on Hydroperoxide Enzyme Levels and the Cardiotoxic Effects of Adriamycin Administration in Mice. Doctoral dissertation, The Ohio State University, 1984.

38. Kanter, M.M., Hamlin, R.L., Unverferth, D.V., Davis, H.W., Merola, A.J. Effect of Exercise Training on Antioxidant Enzymes and Cardiotoxicity of Doxorubicin. Journal of Applied Physiology, Vol. 59, no. 4, pp. 1298-1303, 1985.

39. Katz, L.D., Morton, G.G., Raporort, S., Ferrannini, E., DeFronzo, R.A. Splanchinc and Peripheral Disposal of Oral Glucose in Man. Diabetes, Vol. 32, pp. 675-679, 1983.

40. Keele, B.B., McCord, J., Fridovich, I. Further Characterization of Bovine Superoxide Dismutase and its Isolation from Bovine Heart. Journal of Biological Chemistry, Vol. 246, p. 287, 1971.

41. Kochan, R.G., Lamb, D.R., Reimann, E.M., Schlender, K.K. Modified Assays to Detect Activation of Glycogen Synthase Following Exercise, American Journal of Physiology (Endocrinological Physiology), Vol. 240, pp. E197-E202, 1981.

42. Leier, C.V., Huss, P., Lewis, R.P., et al. Drug-induced Conditioning in Congestive Heart Failure. Circulation, Vol. 65, no. 7, pp. 1382-1387, 1981.

43. Levy, M.N. and Talbot, J.M. Cardiovascular Deconditioning of Space Flight. The Physiologist, Vol. 26, no. 5, pp. 297-303, 1983.

44. Liang, C., Tuttle, R.R., Hood, W.B. et al. Conditioning Effect of Chronic Infusions of Dobutamine. Journal of Clinical Investigations, Vol. 64, pp. 613-691, 1979.

45. Lowry, O.H., Passonneau, J.V. Flexible System of Enzymatic Analysis. New York: Academic Press, 1972.

46. Mcclean, D., Fishbein, M.C., Braunwald, E., Maroko, P.R. Long Term Preservation of Ischemic Myocardium After Experimental Coronary Artery Occlusion. Journal of Clinical Investigation, Vol 61, pp. 541-551, 1978.

47. Mahle, C.D., Tejwani, G.A., Hannisian, S.H., et al. The Effect of Long Term Aerobic Exercise on Weight Gain, Food and Water Intake, Federation Proceedings, 1986.

48. Michel, E.L., Rummel, J.A., Sawin, C.F., et al. Results of Skylab Medical Experiment M171-Metabolic Activity. In: Biomedical Results from Skylab, R.S. Johnson and L.F. Dietlein (Eds.) Washington D.C.: NASA, pp. 372-386, 1977.

49. Miller, W.J., Sherman, W.M., Ivy, J.L. Effect of Strength Training on Glucose Tolerance. Medicine and Science in Sports and Exercise.

50. Mondon, C.E., Dolkas, C.B., Reaven, G.M. Effect of Confinement in Small Space Flight Cages on Insulin Sensitivity of Exercise-Trained Rats. Aviation, Space and Environmental Medicine, Vol. 54, 919-922, 1983.

51. Mondon, L.E., Dolkas, L.B., Reaven, G.M. Enhanced Insulin Sensitivity with Exercise Training-How Long Does It Last? Clinical Research, Vol. 31, pp. 94A, 1983.

52. Monier, S., and Y. Le Marchand-Brustel. Effects of Insulin and IGF-I on RNA Synthesis in Isolated Soleus Muscle. Molecular Cellular Endocr., Vol. 37, pp. 109-114, 1984.

53. Morey, E.R. Spaceflight and Bone Turnover: Correction with a New Rat Model of Weightlessness. Bioscience, Vol. 29, pp. 168-172, 1979.

54. Morey, E.R., Sabelman, E.E., Turner, R.T., Baylink, D.J. A New Rat Model Simulating Some Aspects of Space Flight, The Physiologist, Suppl. pp. 523-524, 1979.

55. Musacchia, X.J., Steffen, J.M., Deavers, D.R. Rat Hindlimb Muscle Responses to Suspension Hypokinesia/Hypodynamia, Aviation, Space and Environmental Medicine, Vol. 54, pp. 1015-1020, 1983.

56. Newsholme, E.A., Leech, A.R. Oxygen toxicity and the superoxide radical. In: Biochemistry for the Medical Sciences. John Wiley and Sons, New York, pp. 152-159, 1984.

57. Oganov, V.S., Skuratova, S.A., Potapov, A.N., Shirvinskaya, M.A. Physiological Mechanisms of Adaptation of Rat Skeletal Muscles to Weightlessness and Similar Functional Requirements. The Physiologist, Vol. 23, no. 6, Suppl, pp. S16-S21, 1980.

58. Oganov, V.S. Results of Biosatellite Studies of Gravity-Dependent Changes in the Musculo-Skeletal System of Mammals. Physiologist, Vol. 23 (Suppl.), pp. S16-S21, 1980.

59. Pagani, E.D., R.J. Solaro. Swimming Exercise, Thyroid State, and the Distribution of Myosin Isoenzymes in Rat Heart. American Journal of Physiology, no. 245 (Heart Circulation Physiology, Vol. 14), pp. H713-H720, 1983.

60. Popovic, V. Antiorthostatic Hypokinesia and Circulation in the Rat. The Physiologist, Vol. 24, pp. 515-516, 1981.

61. Rudy, Y. Plonsey, R., Liebman, R. The Effect of Variations in Conductivity and Geometrical Parameters on the Electrocardiogram, Using an Electric Spheres Model. Circulation Research, 1979.

62. Saere, P.A. Enzymatic Method for Quantification of Citrate Synthase, Methods of Enzymology, Vol. B, p. 3, 1969.

63. Sandler, H. Effects of Bedrest and Weightlessness on the Heart. In: Hearts and Heart-Like Organs. G.H. Bourne, Ed. New York: Academic, Vol. 2, pp. 435-523, 1980.

64. Schaible, T.F., Malhotra, A., Clambrone, G.J., Scheuer, J. Chronic Swimming Reverses Cardiac Dysfunction and Myosin Abnormalities in Hypertensive Rats. Journal of Applied Physiology, Vol. 60, pp. - , 19 .

65. Schalch, D.S., and T.C. Cree. Protein Utilization in Growth: Effect of Calorie Deficiency on Serum Growth Hormone, Somatomedins, Total Thyroxine (T_4) and Triiodothyronine, Free T_4 Index, and Total Corticosterone. Endocrinology, Vol. 117, pp. 2307-2312, 1985.

66. Scharkle, T.F., Scheuer, J. Effects of Physical Training by Running or Swimming on Ventricular Performance of Rat Hearts. Journal of Applied Physiology, Vol. 46, pp. 854-860, 1979.

67. Schoenle, E., J. Zapf, C. Hauri, T. Steiner, and E.R. Foresch. Comparison of in Vivo Effects of Insulin-Like Growth Factors I and II and of Growth Hormone in Hypophysectomized Rats. Acta Endocrinol., Vol. 108, pp. 167-174, 1985.

68. Shellock, F.G., Swan, H.J.C., Rubin, S.A. Early Central Venous Pressure Changes in the Rat During Two Different Levels of Head-Down Suspension. Aviation, Space and Environmental Medicine, Vol. 56, pp. 791-795, 1985.

70. Simard, C., M. Lacaille, and J. Vallieres. Enzymatic Adaptations to Suspension Hypokinesia in Skeletal Muscle of Young and Old Rats. Mechanisms of Ageing and Development, Vol. 33, pp. 1-9, 1985.

71. Smith, A.T., D.R. Clemmons, L.E., Sturek, D.R. Lamb, and A.C. Snyder. Somatomedin-like Activity and Muscle Hypertrophy. IRCS Med. Sci. Physiol., Vol. 9, p. 760, 1981.

72. Sullivan, M.J., Merola, A.J., Timmerman, A.P., Unverferth, D.V., Leier, C.V. Drug-Induced Aerobic-Enzyme Activity of Human Skeletal Muscle During Bedrest Deconditioning. Journal of Cardiopulmonary Rehabilitation, Vol. 6. pp. 232-2378, 1986.

73. Thomas, J.A., Schlender, K.K., Larner, J. A Rapid Filter Paper Assay for UDP-Glucose-Glycogen Glucosyntransferase, Including an Improved Biosynthesis of [UDP-^{14}C] - glucose. Analytical Biochemistry, Vol. 25, pp. 486-499, 1968.

74. Terjung, R.L. Muscle Fiber Involvement During Training of Various Intensities. American Journal of Physiology, Vol. 230, pp. 946-950, 1976.

75. Wronski, T.J., Morey-Holton, E.R. Skeletal Response to Simulated Weightlessness: a Comparison of Suspension Techniques. Aviation Space and Environmental Medicine, Vol. 58, pp. 58-63, 1987.

Appendices can be obtained from

Universal Energy Systems, Inc.

FINAL REPORT

AN INVESTIGATION OF COMPUTER COMMUNICATIONS USING KNOWLEDGE-BASED SYSTEMS

Sponsored by
Air Force Office of Scientific Research
Bolling AFB, DC
Contract No. F49620-85-C-0013/SB5851-0360
Universal Energy Systems Project 760

Submitted by
Gerald R. Graves, Ph.D.-Principal Investigator
John M. Usher-Research Associate

December 1987

Department of Industrial Engineering
Louisiana State University
Baton Rouge, Louisiana 70803
(504) 388-5112

ABSTRACT

"An Investigation of Computer Communications Using Knowledge-Based Systems" is sponsored by the Air Force Office of Scientific Services under Contract No. F49620-85-C-0013/SB5851-0360, and administered under Universal Energy Systems Project 760. The research investigates the use of a knowledge-based system for performing some of the functions associated with communications between heterogeneous computer systems. The Manufacturing Message Format Standard (MMFS) of the Manufacturing Automation Protocol is further investigated. A prototype knowledge-based system which translates between MMFS and a programmable logic controller and a robot controller is presented. The prototype is written in Prolog and implemented on an IBM/AT.

FORWORD

This report describes work completed under Contract Number F49620-85-C-0013/SB5851-0360 during the period from January 1 to December 31, 1987. This effort investigates the potential for applying a knowledge-based systems approach to computer communications.

The project is sponsored by the Air Force Office of Scientific Services, Bolling AFB, DC, administered by Universal Energy Systems, Dayton, Ohio, and conducted by researchers in the Department of Industrial Engineering at Louisiana State University in Baton Rouge, Louisiana. Major Steven LeClair of the AFWAL Materials Laboratory at Dayton, Ohio is the principal Air Force representative, and Dr. Gerald R. Graves and Mr. John Usher are the principal Louisiana State University researchers.

The report is divided into four chapters and two appendicies. The chapters include:

I. Introduction
II. Potential Applications of Knowledge-Based Systems in Computer Communications
III. A Knowledge-Based Systems Implementation of MMFS
IV. Concluding Remarks

The appendicies include a program listing and operating instructions.

CHAPTER I

INTRODUCTION

1.1 Purpose of the Research

This research project is concerned with communications in a heterogeneous computer system environment. The purpose of the effort is to investigate the use of knowledge-based systems (a broader classification which contains expert systems) in providing for computer communication functions. The intent is to complement ongoing efforts aimed at establishing communication protocol specifications for computer integrated manufacturing (CIM).

1.2 Background

For successful implementation of CIM an effective, efficient, and robust integrated computer system must be available. The integrated system must be effective, ensuring that communications between application processes are relatively error-free, and that the communication system is reliable and transparent to users. The integrated system must be efficient, providing for the timely transmission of messages and the high utilization of resources. The system must be robust, allowing for a variety of message contents and for a variety of attached processors and networks, and providing expansion capability.

Local area networks (LANs) are the principal components of an integrated system in manufacturing. The nature of a LAN is determined primarily by three factors: the transmission medium, the topology, and the communication protocol. The transmission medium provides the physical connection between processors (e.g., twisted pair, coaxial cable, or optical fiber). The topology details the physical and logical arrangement of the processors (e.g., star, ring, or bus). The communication protocol is a set of rules governing the exchange of data between entities forming the network, and is the focus of this research.

1.2.1 The OSI Model

A computer communication architecture is a structured set of protocols that implement the communication functions. Key elements of these protocols are the syntax (e.g., data format and signal levels), the semantics (e.g., control information for coordination and error handling), and the timing (e.g., speed matching and sequencing). With the rapid introduction of a variety of computer processors and devices into the manufacturing environment a number of communication protocols were developed by computer system vendors. Integrating these heterogeneous processors has proven to be a formidable problem. The International Standards Organization (ISO) attacked this problem by developing a reference model to serve as a basis for protocols. This model, termed Open Systems Interconnection (OSI), presents standards for the exchange of information among systems that are "open" to one

another by virtue of their mutual use of the applicable standard. The OSI model has received widespread acceptance, with communication system vendors describing their protocols in related terms.

The OSI model uses a layered approach in structuring communication functions. The layering principle and operation of the OSI model is depicted in Figure 1. The communication functions are described within seven layers, with the top layer interacting with the user program and the bottom interacting with the physical media. As shown in the figure, when the application program X has a message to transmit across the network to application program Y it transfers the data to a program which resides in the Application layer of the model. Here, certain communication functions are performed, and the data is then passed to the next lower layer. This continues down through the remaining layers, after which the message is transmitted. Upon arrival at the destination system the reverse process occurs. Protocols implement the communication functions by appending header and trailer information to the data. The OSI model specifies what communication functions are to be provided, and a

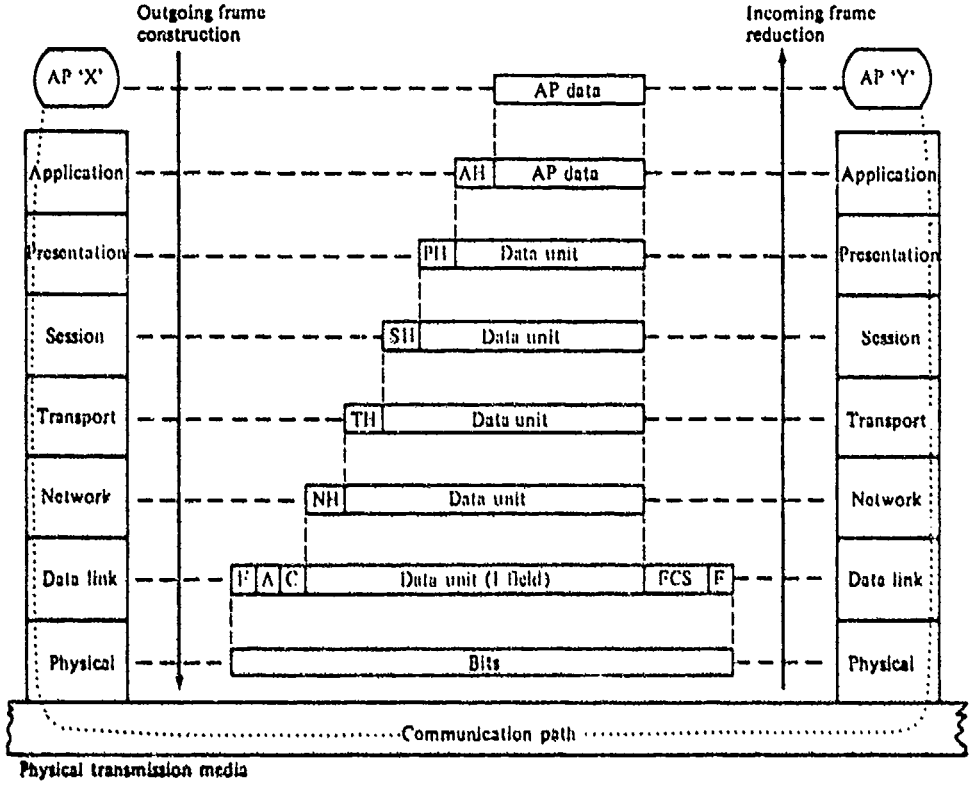


Figure 1.1: OSI Operation. Source: Stallings (5)

layered architecture for providing them. Briefly, the basic functions provided at each layer include:

Physical layer:	physical interface between devices and rules for the transfer of bits
Data Link layer:	error detection and control
Network layer:	establishing, maintaining, and terminating connections
Transport layer:	transparent exchange of data between processes in different system, including error detection and sequencing
Session layer:	dialogue control between applications
Presentation layer:	generic representation of data syntax
Application layer:	access to the OSI environment, including such functions as file transfer and messaging languages

Numerous protocol services are included to provide the functions within each layer. It is important to note that, although the OSI model specifies standard functions to be provided, it does not attempt to specify how these are to be implemented. Thus, different protocol implementations may conform to the OSI model but not be compatible.

This brief description is deceptive. The scope and complexity of the OSI model and standards effort are enormous. A May 1986 listing of the ISO working papers and draft standards addressing the OSI model contains over 100 entries available at a cost of over $2,000. Furthermore, efforts continue to further refine the OSI model.

1.2.2 MAP

In response to the rise in the number of programmable devices and computer systems in use on the factory floor and the need to integrate these, General Motors established a task force in 1980 to develop a solution to their communication problems. In 1982 this task force proposed the development of a non-proprietary communication protocol architecture to be based on the OSI reference model. This architecture, termed the Manufacturing Automation Protocol (MAP), is constrained to communications on the factory floor. A similar effort, termed the Technical Office Protocol (TOP), was initiated by Boeing Computer Services. These efforts have been combined into the MAP/TOP User's Group which now consists of over 300 companies under the administration of the Society of Manufacturing Engineers.

MAP specifies a set of protocols for implementing communications using the OSI 7-layer architecture. For each layer MAP specifies protocols to implement the functions of the OSI model.

Particular attention was focussed on the migration from earlier to later versions of the MAP specification. The intent is to encourage implementation of the latest version without concern for future modifications. The most recent version, MAP 3.0, was released in the summer of 1987. A multi-vendor demonstration of a MAP 3.0 network is planned for the summer of 1988.

In addition to protocols, MAP addresses other issues which are critical to successful integration on the factory floor. Specifications are presented for devices to connect with non-MAP networks and subnetworks, such as bridges, gateways, and routers. The functions of network management are specified, including such services as monitoring, control, configuration, and problem detection and recovery. Also, a User Control Interface is specified which provides services to attached processors allowing for communication with other end users of a MAP system. Lastly, a modified MAP architecture, the MAP Enhanced Performance Architecture, is specified for use with real-time applications.

Although the progress of MAP efforts is encouraging, general acceptance and implementation of MAP are still distant. Consequently, research into techniques for improving computer communications is appropriate.

1.3 Project Chronology

The initial approach taken by the LSU researchers involved investigating the use of a knowledge-based system for communication between heterogeneous computer processors. Specific services to be provided by the knowledge-based system were those associated with the upper layers of the OSI model: the session, presentation, and application layers.

The research effort officially commenced in January of 1987 and involved an analysis of the existing MAP and OSI specifications. This was accomplished through literature review and contact with experts in the field of protocols. It quickly became obvious that almost no research of this nature had been conducted.

Three potential areas for the application of knowledge-based systems were identified, including the Manufacturing Messaging Format Standard (MMFS), the User Control Interface, and network management.

In April it was decided to concentrate on the MMFS for further investigation. Thus, the development of a knowledge-based implementation of MMFS began. Efforts to improve understanding of protocols and to investigate other potential applications of knowledge-based systems in shop floor communications were, however, continued.

The prototype was completed in November of 1987 using Prolog on an IBM/AT. The prototype is designed to interpret between the MMFS and a device-specific format for either a Honeywell programmable logic controller or an NND robot controller. The remaining duration of the project was devoted to analyzing the prototype.

1.4 Scope of the Report

The substance of this report is contained in two chapters. Chapter II discusses potential applications of knowledge-based systems in computer communications, constrained to a MAP environment. Chapter III concentrates on the prototype. After describing in detail the MMFS, the Prolog implementation is presented. This description includes an example and a discussion of the advantages and shortcomings evident in the prototype.

The last chapter of the report contains concluding remarks possible further research. A program listing and instructions for running the prototype are included as appendicies.

CHAPTER II

POTENTIAL APPLICATIONS OF KNOWLEDGE-BASED SYSTEMS IN COMPUTER COMMUNICATIONS

2.1 Introduction

A short-term goal of the research was to identify potential applications for knowledge-based systems in computer communications. To accomplish this goal the question was posed "Why use knowledge-based systems in computer communications?" It would seem that developing standards (such as MAP) would be more desirable, especially in light of the advances made possible through other standardization (e.g., the ASCII code and the RS232 interface). However, it is extremely difficult to establish a rigorous standard which provides the desired flexibility. This is especially true in computer communications which requires a broad range of services with a variety of options to accommodate a myriad of systems, devices, and applications. Yet, flexibility is an obvious necessity in the emerging field of computer integrated manufacturing. Rather than a rigid standard with minimal flexibility, especially at the higher layers of a protocol architecture, it would seem that a knowledge-based system might enhance protocol standards by allowing for expanded options and negotiation between communicating entities.

With this theme in mind, then, the process of identifying candidate functions was conducted. Specifically, those activities included in MAP which allow for negotiation, optional services, or some form of decision-making were considered. Three applications were identified as potential candidates for further study: the User Control Interface, Network Management, and the Manufacturing Message Format Standard.

2.2 The User Control Interface

Generally, two kinds of applications communicate within a MAP system- those executing within a MAP node, and those residing in an attached processor. The User Control Interface (UCI) provides services to attached processors by exposing a transport layer interface. Services provided by the UCI include: establishing/monitoring/closing connections to other MAP stations; sending data and expedited data; signalling message arrivals to inactive queues; and indicating errors in date units.

Through the UCI, then, users have the flexibility to use non-MAP protocols for providing upper layer services (presentation, session, application layers). Since a variety of these protocols may exist among end users, and each may have advantages and drawbacks, the opportunity to negotiate protocols at the higher layers exists. Furthermore, the selection of these protocols may change as the operating performance of the network changes. For instance, a protocol implementation which executes rapidly may be utilized during periods of heavy network traffic, with the tradeoff being a degradation in the error detection service.

A knowledge-based system approach might be advantageous in conjunction with the UCI. However, it is believed that an investigation of this application would require a substantial effort, extending well beyond the period of this research project.

2.3 Network Management

Perhaps the most fertile area for the application of knowledge-based systems in computer communications is in network management. The network management function is responsible for ensuring correct operation of the network and reporting on the usage of the network media. The data and programs implementing the network management function may reside at either a central location or at distributed locations. Some of these functions may even be implemented on a separate parallel network. Furthermore, portions of network management are implemented within the protocol layers, complicating this function even more.

Traditionally, network management has included such functions as: fault detection, identification, and recovery; performance management and control; network initialization and reconfiguration; resource management; network access and security; and accounting functions. However, in a MAP environment fundamental changes are anticipated, primarily due to the limitations in the scope of the network and the presence of real-time manufacturing applications.

Initially, MAP requires the implementation of four functions:

1. Configuration Management: determination and control of the state of the system; this is the logical and physical configuration of the system.

2. Performance Management: control and assessment of the performance of the nodes and network operation.

3. Event Processing: generation and interpretation of notification of unsolicited significant occurrences (state changes).

4. Fault management: diagnosis of failures using tests initiated on the network by the manager.

All of these functions lend themselves to knowledge-based systems implementations. However, to complete meaningful research which addresses these potential applications would require more resources (e.g., a dedicated network) and time than was available for this effort.

2.4 The Manufacturing Message Format Standard

MMFS is a messaging language intended to provide a common, versatile mechanism for machine-independent process-to-process information exchange. This standard defines the form and meaning of messages and their components, including the syntax and grammar of a message, and the semantics, or vocabulary of standard

fields in a message. The purpose is to ensure that messages from one application can be processes by another, and that standard meanings for the components of a message ensure a common interpretation of a message by both the sender and the receiver.

The features desired for MMFS (e.g., flexibility, efficiency, ease of expansion) encouraged further exploration of this standard. Furthermore, it was felt that the development of a prototype implementation represented a reasonable effort for this project. In the next chapter additional details regarding MMFS are presented, along with the research prototype implementation.

3.5 Remarks

In order to devote adequate time to development of the knowledge-based prototype it was necessary to limit investigatory efforts addressing potential knowledge-based systems applications in communications. Consequently, the preceding discussion is not intended to be a comprehensive review. Certainly, other candidate applications exists, but those presented are the more obvious ones.

CHAPTER III

A KNOWLEDGE-BASED SYSTEMS IMPLEMENTATION OF MMFS

3.1 The Manufacturing Message Format Standard

As stated in the Manufacturing Automation Protocol (MAP) document, the "...purpose of the Manufacturing Message Format Standard (MMFS) is to facilitate the transfer of digitally encoded information between various, possibly disparate, devices" [1]. The translation facility for the MMFS messaging language resides in the application layer (layer 7) of the OSI architecture. The MMFS standard provides a description of the syntax and semantics of MMFS messages. The use of a standard for network messaging ensures that all attached devices will have the capability to transmit and receive recognizable messages. The MMFS was design based on a number of important properties which include [1]:

- providing a syntax which is consistent, efficient, and flexible,
- easy expansion of the standard's functionality,
- minimizing the need for reprogramming, and
- allowing for implementation of only those features needed for a given application.

The following discussion is meant to provide an overview of the fundamentals of the MMFS to aid in understanding the prototype implementation of the knowledge-based interface to be presented later.

3.1.1 MMFS Syntax

The MMFS is composed of two principle building blocks, the "field" and the "data stream". The field itself contains two subfields. The first is an identification subfield providing a label for the field and indicating the type of information contained in the content subfield, the second subfield. The content subfield carries the data of the field and is limited to a length of up to 127 octets (bytes). The data stream is similar to the field in that it too carries data, but it does not provide any label as to the type of data it contains. The data stream can also hold a larger quantity of data (10**300 octets) than the content subfield. The data stream would be employed when "it would be awkward and inefficient to place the data in fields," and/or there is the desire "to send application specific data in a non-standard format" [1].

The MMFS uses the construct of a "group" which is composed of sets of fields. Nesting of groups within groups is allowed with the concept that the field is the primary group where no nesting is allowed. The use of groups allows for delimiting of sets of fields within a message. Delimiting of the groups is provided by the leading field of each group referred to as a "grouper". In the MMFS, there exists five type of groupers, the

grouper, the concatenation grouper, and the data stream grouper. A message to be transmitted is composed of an outermost group containing all groups, fields, and data streams within it [1]. Each message must begin with a grouper field which will denote, by some means, the length of the message. This starting field must contain either the octet-count, group-count, or parenthesis grouper.

The subfields of the field can be specified in either of two formats. The first is a short format requiring one byte. This format is denoted by the presence of a zero in the most significant bit of that octet. The long (or multi-octet) format uses the first octet as a length octet to specify the number of bytes pertaining to that subfield which follow. The multi-octet format is detected by a one in the most significant bit position of the first octet. This results in a maximum possible subfield length of 128 octets (one octet to specify the length and a maximum of 127 octets which follow).

3.1.2 Notation

The MMFS notation identifies a field by the use of a single set of angle brackets to enclose the field. Similarly a data stream is denoted by the use of double angle brackets to set off its contents. To display the two octets of a field, the subfields are separated by the use of a colon. When displaying a multi-octet subfield, the first octet (the length octet) is separated from the rest by use of a semicolon. Examples of these rules are:

notation for a field	< field >
notation for a data stream	<< data stream >>
short format subfields	< ID : content >
multi-octet content subfield	< ID : length octet ; data >

These rules of notation will be used to display the MMFS messages for the examples later in this paper.

3.1.3 Semantics

The semantical elements are the actual fields employed in an MMFS message, and their meanings [1]. For the convenience of discussing the MMFS fields, mnemonics are assigned for each of the fields. These mnemonics will be used to write the message in a readable form, but will not be transmitted as part of the message. The binary value of the hexadecimal equivalent of the mnemonic is what will be transmitted.

Fields are represented by two and three character mnemonics. The two character mnemonics are used to represent a class of fields which have the same identification subfield. The content subfield of the fields with two character mnemonics can carry data (i.e. numeric, string, untyped, etc.) or it can contain a function code. The function codes represent types of the data and all function codes within a class have the same identification subfield. These fields which contain function codes are

identified by the use of a three character mnemonic. Each MMFS content field, which carries data, will have an assigned default data type (e.g. unsigned integer, untyped data, character, etc.). The MMFS fields that were implemented in the prototype knowledge-based interface appear in Table 1.

TABLE 1

Implemented MMFS fields

Mnemonics	ID Subfield	Content Subfield	Explanation
OC	Ø1	UI	Octet count grouper
DS	Ø5	UI	Data stream grouper
FA	Ø7	UD	Fully qualified address
CT	ØB	UI	Count
CH	19	Data	ASCII character
TN	1F	Data	Transaction number
PC	2Ø	codes:	Procedural control
CRQ		Ø1	Cmd/Req with response
FRS		Ø3	Final response
I2	41	codes:	PLC instructions
REA		Ø1	Read
WRI		Ø2	Write
RUL		Ø4	Upload read
WDL		Ø9	Download write

The fields will be described below in reference to an example message used in the prototype implementation. The example appears in the application of the interface to provide for message translation between an MMFS host and a programmable logic controller (PLC). The following MMFS message sent to the device informs the PLC that it is performing a download of 2 program instructions to be stored starting at memory location Ø9Ø8 hex. This message in mnemonic and hexadecimal form is:

```
<ØC:ØC><TN:Ø1><CRQ><WDL><FA:Len=2;Ø9Ø8><DS:Ø6><<112233445566>>
 Ø1ØC   1FØ1   2ØØ1 41Ø9  Ø782    Ø9Ø8  Ø5Ø6    112233445566
```

The message begins with an octet count grouper specifying that there are 12 octets in the message (does not include the data stream). This is followed by the transaction number field which labels each transaction with a number for coordinating command requests with their associated response (which carry the same transaction number value). The transaction number is always followed by a procedural control field. In this example, a command/request with response, CRQ, has been issued indicating

that the receiving entity is to perform the required action and send some type of response back (i.e. acknowledgement, error code, etc.). The FRS response field, appearing in Table 1, would be the procedural control field in the response message. This field indicates a positive final response to the request.

The WDL field is from "instructional group 2" of the MMFS which caters to the PLC. As the table indicates, WDL is the mnemonic for a request to download data to the PLC. The FA field that follows is one possible field used for data definition. This field allows for specification of a fully qualified address which is machine dependent. This provides an address to indicate where data is to be read from or, in this example, written to. The count field, CT, is also a data definition field and is used to indicate the number of data items being referred to by the other immediate fields. For example, in communication with the PLC the count field can be used to indicate the number of instructions which are to uploaded from the device, beginning at the address given by the FA field. The FA field in the above example message uses a multi-octet format for the content subfield. This subfield indicate that the address can be found in a two byte segment following the length octet (i.e. Ø908).

The next field in the example is the data stream grouper. The content subfield of the DS field specifies the number of octets that will appear in the data stream immediately following the DS grouper. The data in this example is considered to be untyped data and will be interpreted by the receiving device in the form the device expects. Therefore, since no data type was specified it is the job of the sender to ensure that the data type of the data appearing in the data stream match what is expected by the receiving entity.

The example presented here is one of six that appear in the appendix. The reader is referred to those other examples to examine more of the constructs involved in MMFS messaging. There exist many more MMFS fields than those discussed here. These can be found in Appendix 6 of the MAP document [1].

3.2 Implementation of the Knowledge-Based Interface

As was described earlier, the knowledge-based interface provides for the translation of messaging formats between each of the disparate formats used by the attached devices. To simplify this implementation, operation of the prototype interface will be limited to performing message translation between only two different messaging formats.

One messaging format is assumed to be the standard format used for all communications on the underlying network. For the prototype, the Manufacturing Message Format Standard (MMFS) was selected to be the network standard, because of its designation as the standard messaging language for the Manufacturing Automation Protocol (MAP) specification. The second messaging language is that of a device attached to a network node, and will be dependent on the device manufacturer. Therefore, operation of the interface will result in the translation of all messages passing between an attached device, using a device-specific format (DSF), and the underlying network, using the MMFS format.

to the network which do not conform to the network messaging standard. Instead of requiring that the user design a custom interface for each non-standard device, they would build a knowledge base for each device using the appropriate information obtained from the device's manuals. This task requires less time and doesn't demand a skilled programmer to perform the task.

The prototype implementation, as it pertains to the two examples cited, will be discussed in more detail in the sections that follow. The reader is reminded that this initial implementation is meant only to demonstrate the potential of using such a design for message translation.

3.2.1 Prototype Applications

The prototype was developed for two different applications. One application involved the connection of a Honeywell IPC-620 programmable controller (PLC) as the attached device. The knowledge-based implementation allows for the use of four of the available PLC instructions. These instructions are the download instruction to provide for transfer of a program from a host to the PLC, an upload instruction allowing for the host to gain access to the PLC program, and instructions to allow for reading and writing to the PLC registers [2]. The PLC was selected because it utilizes integer data in a structured messaging format with each message field explicitly defined in terms of its length and possible values.

The second example deals with connecting an NND Assembly Robot 88. For this example, only two instructions were implemented. The first is a command for the robot to return to its origin, and the second is a command to upload the contents of the robot memory to the host (network)[3]. The robot was selected because it demonstrated a different data type and format from the PLC. The robot uses a character data message with no structured format beyond the use of blanks and commas as separators. The end of a message is determined by the use of two indicator bytes.

These two extremes will help demonstrate the applicability of the knowledge-based interface to widely varying devices in terms of their required message formats. The actual program code for the control program and each of the example knowledge bases appears in the appendix.

3.2.2 Programming Language

The prototype was developed using an IBM PC/AT class of microcomputer. The implementation of the knowledge-based system required that the system be developed from scratch because the nature of the tasks it was to perform didn't fit the capabilities of the knowledge-based shells currently available.

The prototype program was written using Turbo Prolog, a Borland Corp. product, because this language allows for quick development of the program and provides facilities for simplified debugging. In addition, Turbo Prolog produces compiled code which can be executed on any IBM-compatible microcomputer without requiring the Prolog compiler program.

Since Prolog operation is based on the use of a stack, the operation of the prototype is limited by the stack space available. This problem is only a software limitation and can be overcome by improvements in the programming techniques used, or better yet, by the use of a different programming language such as LISP or C.

This implementation does not demonstrate the programming techniques required for such an interface. This prototype is meant only to demonstrate the feasibility and usefulness of such a design for a knowledge-based interface. In all likelihood, a production system would have to be ported to a lower level language such as LISP or even the more portable and popular "C" language. Important factors will be the speed of the execution and the ability to generate a DSF knowledge base that can be easily constructed by an semi-experienced communications engineer or technician.

3.2.3 Program Description

For implementation of the interface, a knowledge-based program design was chosen. Such a design allows for the separation of the device dependent knowledge from the routines needed to carry out the translation and transmission of the messages (i.e. the control program). The knowledge-based interface can be viewed, as shown in Figure 3.1, to be composed of three main components, the control program, the device-specific knowledge base (DSF) and the MMFS knowledge base. The control program has the responsibility of performing the translation, and receiving and transmitting the messages to the appropriate connected device (i.e. network or device). To perform the translation, the control program requires knowledge concerning the messaging format of the devices attached. This information is contained in the DSF and MMFS knowledge bases.

Program Structure. The Turbo Prolog program, as depicted in the block diagram of Figure 3.2, consists of program declarations and program clauses. In Turbo Prolog, the user is required to make several declarations. The first concerns the variables that will be used in the working memory of the system (often referred to as the dynamic database). These variables will hold the key information needed to perform the translation and construction of the new formatted message. The variable values stored there will also be used to hold information used to construct the appropriate response. Figure 3.3 presents the contents of working memory, categorized as the incoming message, MMFS-specific elements, and device-specific elements.

In addition to the database variables, each of the clauses used in the program must be declared along with the objects of the statement and their domain type (i.e. string, real, integer, etc.). The next section of the program is the heart of the Prolog program and contains all the program clauses and their components. These clauses can be further subdivided into those pertaining to the control program and those of the knowledge bases.

Control Program. A further breakdown of the control program

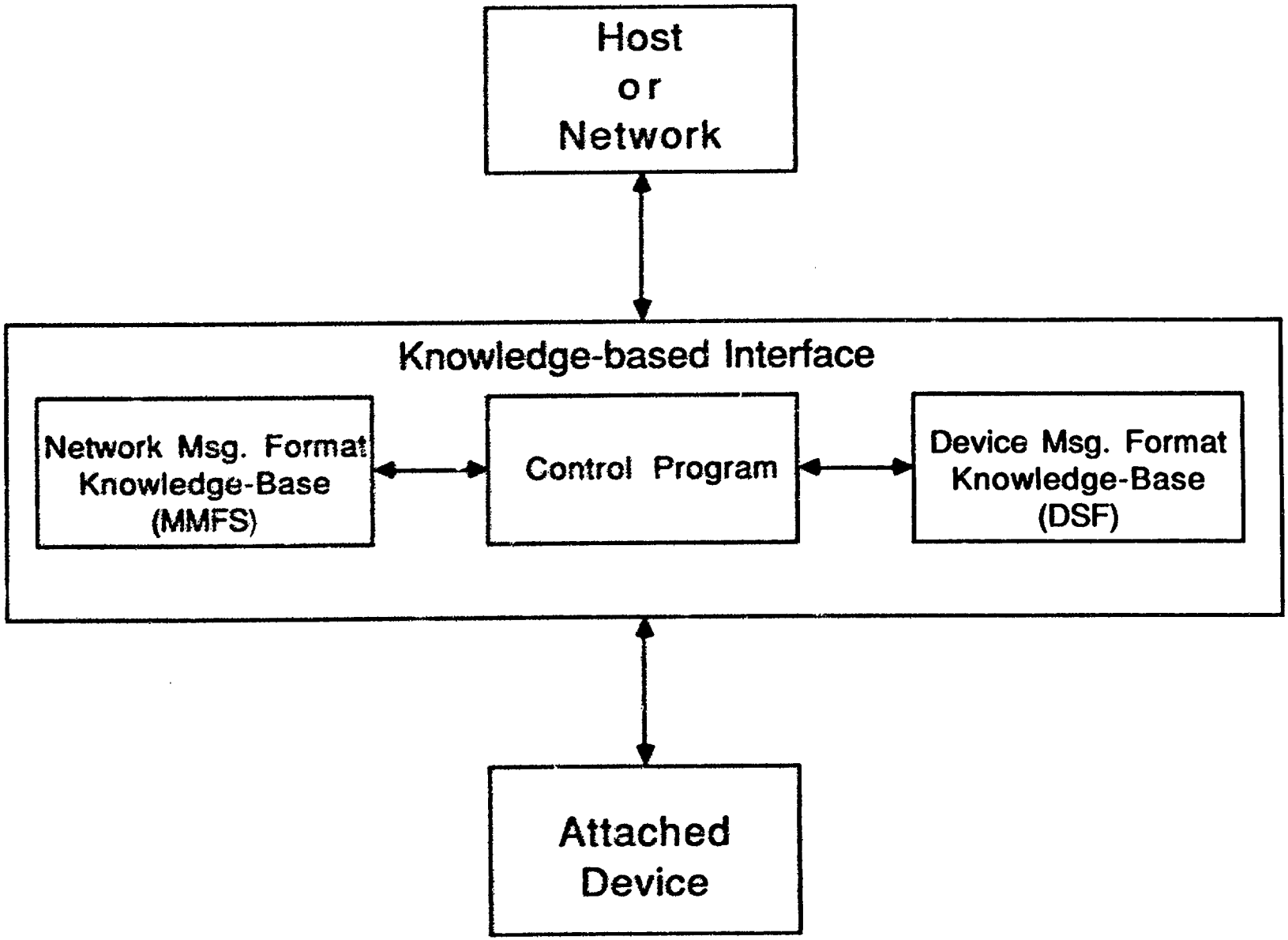


Figure 3.1: Knowledge-Based Messaging Inteface

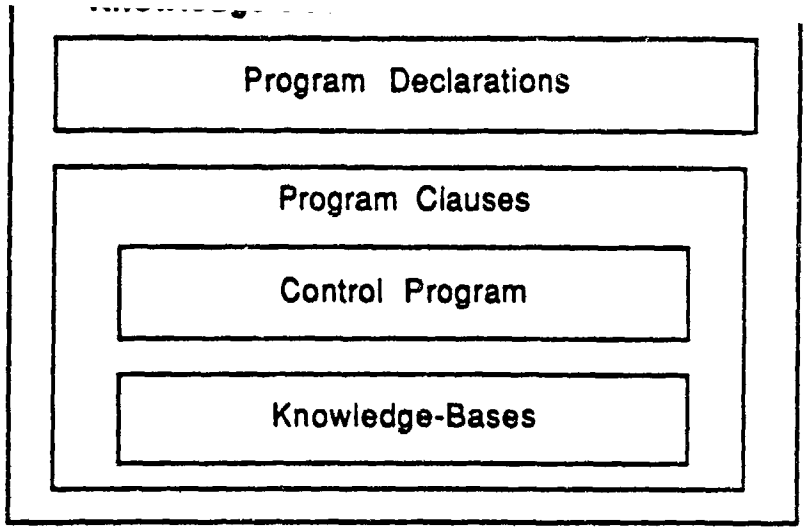


Figure 3.2: Block Diagram of Program

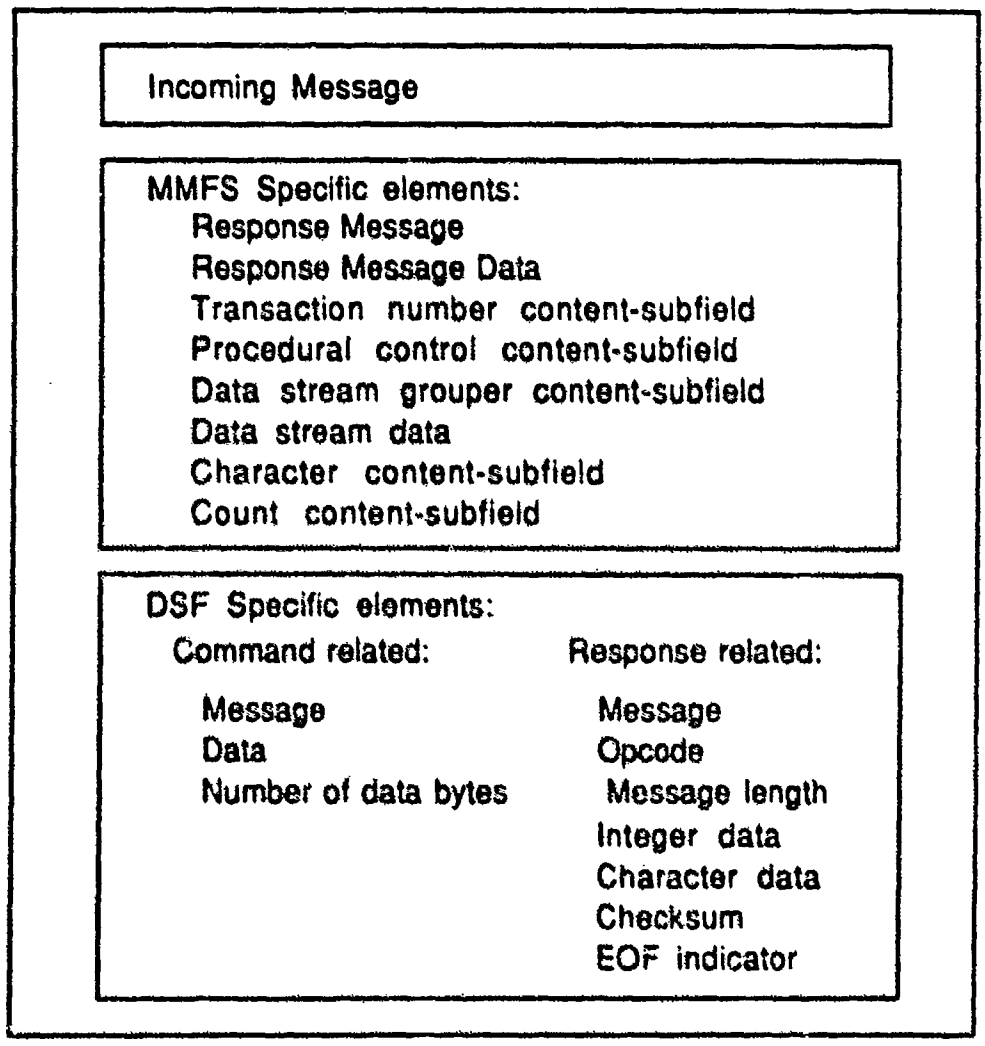


Figure 3.3: Contents of Working Memory.

is given in Figure 3.4. This figure shows that the control program is composed of four main modules. This grouping of the clauses into modules was performed to allow for ease in finding the appropriate routine and to aid in exposing the position of the required functions. The first module houses the main routines of the program. These are the clauses that call the lower level routines which perform the functions of the system. Included in this module are clauses which coordinate functions such as initialization the system, awaiting the receipt of a message, parsing of a message, assembling of a message, transmission of a message, and the clearing of working memory.

The next two modules contain the routines which perform the services the MMFS and DSF messaging standards, respectively. These routines use the knowledge housed in the corresponding knowledge bases. The last module of the control program, houses the basic system primitives. These are functions required by the system which are not specific to this interface and can be considered common programming functions not provided by the Prolog implementation (e.g. converting variables from string to real, real to integer, hex to string, etc.).

Some specific characteristics of the MMFS messaging language are included in the MMFS-specific module of the control program. This limits the use of this prototype interface to provide for translation between a given device and the MMFS standard only. The same technique could be used to provide for another networking standard in place of MMFS, but would require editing of the control program.

MMFS Knowledge Base. A block diagram of the knowledge base component of the program is shown in Figure 3.5. Two knowledge bases are provided, one for the MMFS messaging standard and one for the DSF messaging standard. Since MMFS is the default standard messaging language for the prototype interface, the knowledge base is not changed for each application. This knowledge base is to be altered only to vary the conformance level of the MMFS implementation. As was discussed earlier, the conformance level of the MMFS implementation determines what fields are allowable for use in messaging. For this prototype only eight of the many possible fields are included. To change the level of conformance of the implementation the user would only be required to add and/or edit the information contained in the MMFS knowledge base. This allows for simplified maintenance of the system from the viewpoint of the network messaging standard.

The parsing predicates in the MMFS knowledge base are used when the system is parsing a MMFS message. As a message is parsed the predicates perform functions appropriate for the identified field and store variable values in working memory. These functions include such tasks as storing the current transaction number, storing the source or destination address, determining the length of the data to follow, etc. The assembly predicates support the assembly of an MMFS message. The current implementation only provides the proper procedural control field to be returned in a MMFS response message.

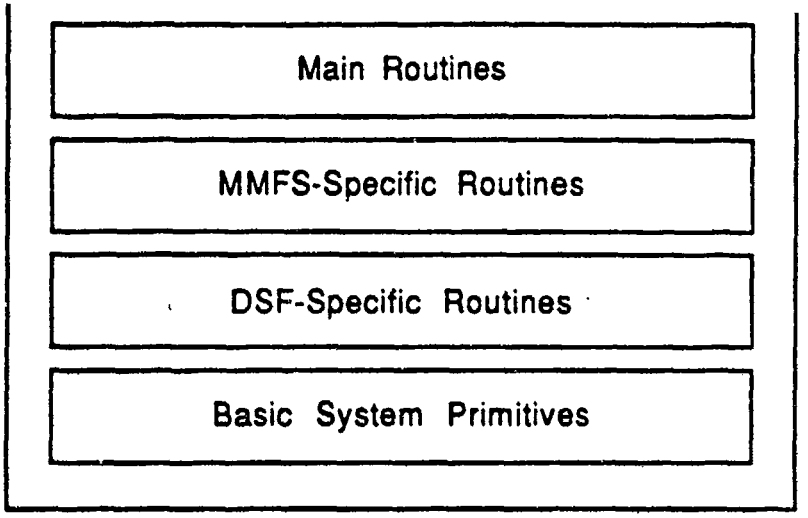


Figure 3.4: Block Diagram of Control Program

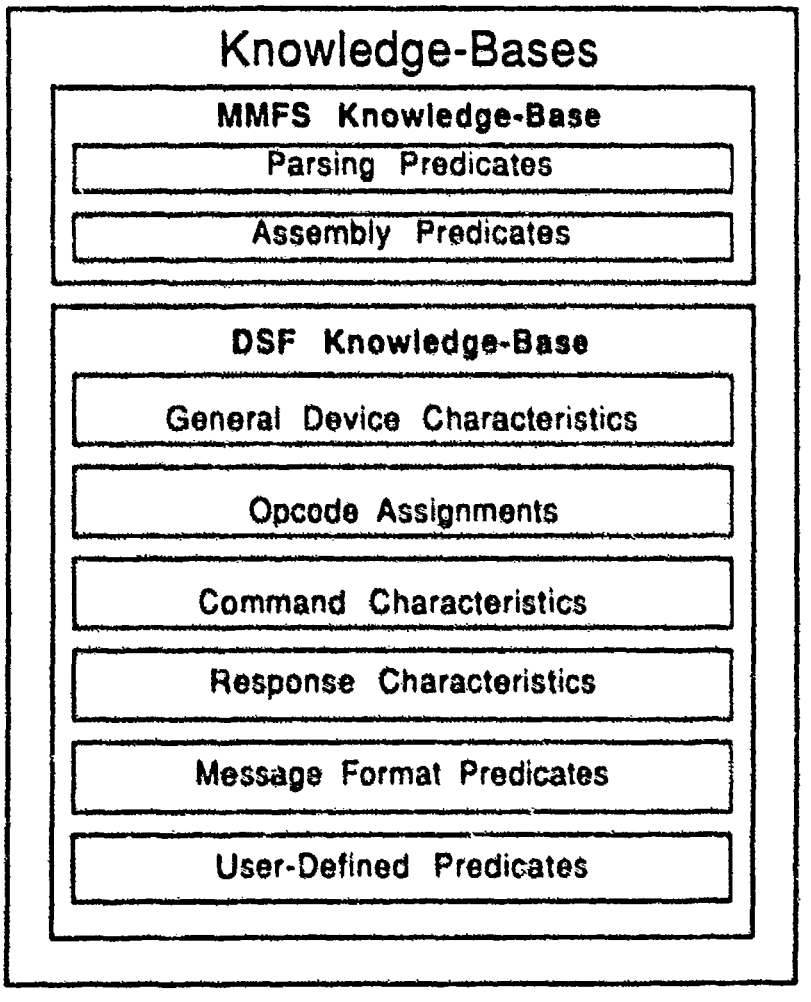


Figure 3.5: Block Diagram of Knowledge-Bases

DSF Knowledge Base. The DSF knowledge base provides for flexibility in the system. This is the component of the system which the user must address in applying the interface to different devices. There are six components of the DSF knowledge base (see Figure 3.5). The first component houses the general device characteristics. These characteristics include such things as whether the device message format is structured or not, the data type of the message, the identity of the header and its length (in bytes), and the identity of the end of transmission element and its length.

The second component directly maps the MMFS instructions (such as those of the PLC instructional group) to the corresponding opcodes of the attached devices, if it has any. The third and fourth components of the DSF knowledge base house the command and response characteristics for the device, respectively. It is here that the message format of the device is specified by providing the information on the presence and length of certain fields for the possible device commands and responses.

The next component of the DSF knowledge base specifies the format of the device message. The device format is specified by the user selecting the appropriate message component statements (i.e. header, opcode, checksum, address, etc.) and placing them in the order specified in the device manual. Each of the message component statements calls a routine which performs the required functions for the message at hand. As an example, consider the assembly of the PLC message from an MMFS command. The PLC message knowledge base specifies the message components of header, opcode, message length, integer message, checksum, and trailer, in the "assemble_dsf" predicate. The control program will fetch the needed information from working memory and determine the correct byte values to place in the DSF message being constructed. Also included in this component of the knowledge base are predicates which allow for manipulation of the data passed as the object of a message (e.g., the program instructions in a download instruction, or the register values in a read register instruction).

The last component of the DSF knowledge base includes predeclared generic predicates which are user-defineable. These user-defined predicates allow for increased flexibility by allowing the user to write routines to perform some function not supplied by the basic system. For example, the PLC places a three byte instruction in a four byte field. The MMFS message will only provide the three bytes of an instruction when downloading. Therefore, the user-defined predicates can be used in combination with the data manipulation predicates to create routines which add the extra zero value byte before transmission to the PLC. This same type of scheme had to be repeated on program uploading, but with the removal of the leading zero before construction of the MMFS data field.

The prototype implementation of the DSF knowledge base is not intended to provide for message translation for any given device. The components and predicates included in this implementation provided the needed functions to carry out the translation for a subset of the instructions of two different devices. As the system is expanded to include several more devices a more robust

knowledge base design can be expected. Information required in the knowledge base should be readily available from the manuals for specific devices. Only a familiarity with communications is required to construct a working knowledge base, assuming a clearly-written manual is available. A user-friendly interface would facilitate building the DSF knowledge base. This would help in separating the user from the unfamiliar structure of the knowledge base, and would provide a simplified query type of environment for knowledge acquisition.

3.2.4 Execution Example

Operation of the prototype program is depicted in Figure 3.6. The program remains in an idle state until it receives a message. It then determines from which user the message was received (i.e. a DSF or MMFS message). This is accomplished by examination of the three bytes out of the first six bytes of the message to determine if they match the MMFS format (i.e. octet count grouper, transaction number field, and procedural control field). In higher levels of conformance the octet count grouper might also be a group count grouper, but this could easily be accommodated by extending the identification fields in the MMFS knowledge base. The appropriate path is then followed based on the format of the message received. Each path executes the same major tasks but the routines to perform these task are different.

The first major task is the parsing of the message. Beginning at the head of a message the semantical components are extracted and stored in working memory for later use. Next, the translated message is assembled. This is performed by calling routines which use the information from the knowledge base and working memory to piece together the message elements as specified by the user in the knowledge base. The translated message is then transmitted to the appropriate user. If no response is expected, the system returns to the initial idle state. If the message requires a response, the system will await the receipt of the response and then repeat the steps for translation of the response message (see Figure 3.6).

In the prototype implementation, the program will not accept a non-MMFS message while in the initial idle state. Although the concept of parsing, assembling, and reading a message are identical for both a command and a response, the actual program steps required differ. Therefore, the prototype currently only supports the acceptance of a MMFS command which translates to a DSF command, followed by a DSF response which is translated back to an MMFS response. Future work could modify this shortcoming.

An example of the application case with the use of the PLC is given in Figure 3.7. In this example, a command message is issued by the MMFS host and received by the interface. The message is:

Msg 1: 010C1F0120014109078209080506112233445566

As was discussed section 3.1.3, each of the groups can be identified by the value in the first byte position. A parsing of this

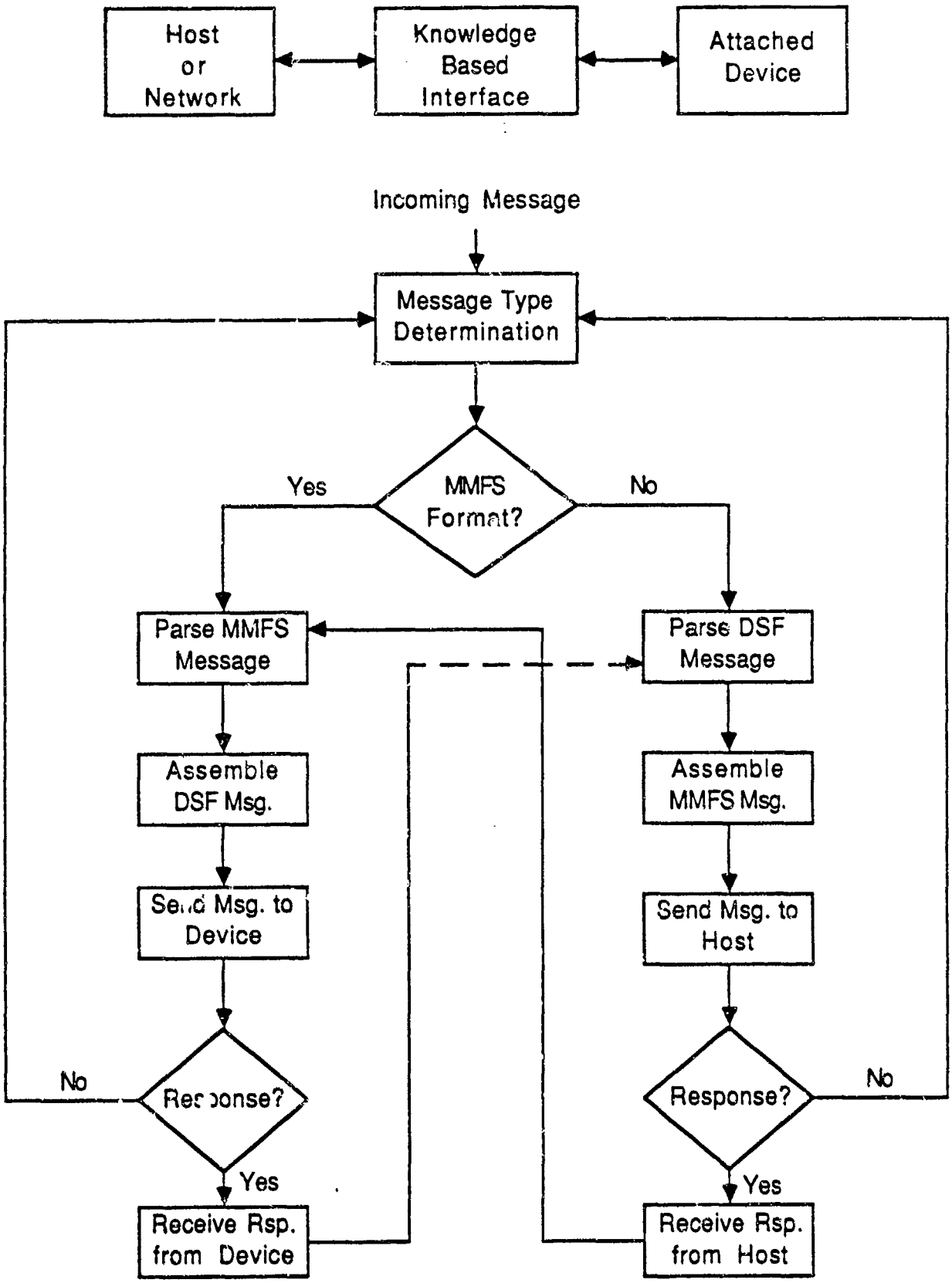
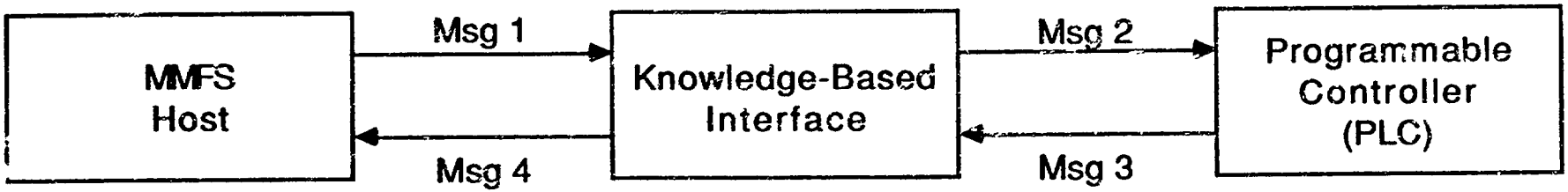


Figure 3.6: Knowledge-Based Messaging Interface.



Msg 1: 010C 1F01 2001 4109 0782 0908 0506 1122 3344 5566

Msg 2: 1,1,1,23,2,36,0,12,0,2,9,8,0,17,34,51,0,68,85,102,195,3

Msg 3: 1,1,130,23,2,0,0,0,27,3

Msg 4: 0106 1F01 2003 4109

Figure 3.7: Messaging Translation Example: Download instruction for the PLC.

message reveals:

```
010C  - octet count grouper - 12 octets follow
1F01  - transaction number is 1
2001  - command request with response required
4109  - a PLC instruction for download write
0782  - signifies a 2 byte fully-qualified address follows
0908  - the address where download will begin
0506  - data stream grouper - 6 octets follow
112233445566 - the data.
```

This information is stored in working memory and the program moves to the next step of assembling the DSF message. The message sent to the PLC is depicted in Figure 3.7 as Msg 2. The commas are not sent but appear only to improve readability. The fields of the PLC message include:

```
header            - 1,1,23,2
opcode            - 36 (download)
message length    - 12
instruction count - 2
address           - 9,8
instruction 1     - 0,17,34,51
instruction 2     - 0,68,85,102
checksum          - 195
end of trans.     - 3
```

This message is received by the PLC and a response (Msg. 3) is given. The response indicates by the appearance of a zero opcode value that no error has occurred. This message is received by the interface and translated into a MMFS response using the values for the transaction number, procedural control type, and instruction type, from working memory. The field value "2003" of the assembled MMFS response (Msg. 4) indicates that the message was received by the device without error and that this is the final response to the previous command. Three other examples for the PLC and two examples for the robot are given in Appendix C.

3.3 Analysis and Conclusions

The design and development of the knowledge-based interface provided enlightenment in a number of areas. First and foremost, the research supports the use of knowledge-based techniques in computer communications. The prototype for MMFS demonstrates the flexibility and expandability that is possible using this approach. This is achieved primarily due to the separation of device-specific knowledge from the control program for translation and transmission of messages. Rather than developing a unique interface program for each device, the knowledge-based interface can be expanded to include device-specific information. Assuming a user-friendly interface to the knowledge base is developed, this expansion could be accomplished without the need for a skilled programmer. Thus, the knowledge-based interface provides increased flexibility.

Furthermore, the design improves expandability by easing the task of modifying the interface to accommodate upgrades in attached devices. This also applies to changes in the conformance level of MMFS. Adding content subfields to working memory and modifying the MMFS knowledge base should provide for this type of expansion.

The design of the knowledge-based messaging interface is modular, with the functions to be accomplished and the information required clearly identifiable. Thus, future research or analysis efforts can benefit from this work, whether the Prolog program is used or one developed in a different language.

Despite the accomplishments made in this study, there are several areas which require further consideration. The first involves a programming problem in the knowledge base. The user is required to specify the message format in three separate predicates. Examination of the PLC knowledge base shows that each of these specifications (see the predicate statements "assemble_dsf_msg", "read_dsf_msg", and "parse_dsf_msg") requires knowledge of the message format. A new predicate is needed which can provide the message format to each of these three routines using a single message format statement, thus streamlining the program.

Now that there is an understanding concerning what the required functions are and where they should be placed in the program, the interface program should be written in a lower level language to explore the execution timing requirements and problems related to the language of implementation.

The use of Prolog for the prototype has been beneficial, but since the execution of the overall program is recursive, the execution time will be limited by the stack space. This is acceptable for research purposes, but could be a problem in some applications in a production environment.

As previously mentioned, the interface program could be extended to allow the interface to handle multiple device message formats. Another area of interest concerns an extension of the interface to encompass the translation between any two device formats, as opposed to one limited to translation between a built-in standard format (such as MMFS for this implementation) and one device format.

Another problem experienced is that the interface as written does not interface with any actual hardware. In its present form, the user is required to enter all the commands and responses for both the device and the MMFS host. This simulation of an implementation was necessary due to hardware problems experienced in providing the lower layer services of communications.

CHAPTER IV

CONCLUDING REMARKS

This research investigated the use of knowledge-based systems in providing computer communications functions. The User Control Interface, network management, and MMFS of MAP were identified as promising candidates for knowledge-based systems implementations. The MMFS was further analyzed and a prototype developed.

Although the prototype is limited in scope, it supports the use of a knowledge-based system for translating between MMFS and device-specific messages. Improvements in flexibility and expandability are possible with this approach. The prototype effectively separates the knowledge base from the control program, resulting in the improvements.

Additional research is needed to thoroughly analyze the knowledge-based system and to convincingly demonstrate its advantages with MMFS. The follow-on research will consist of three phases. First, a third device will be selected and included in the system. Hopefully, minimal changes to the control program will be required. Adding the additional device will reveal any remaining deficiencies in the control program. The second phase will involve translation of the knowledge-based system into the C programming language. This language is more appropriate for system implementation, especially in a real-time manufacturing environment. The third phase will entail the development of a user interface. This is essential for demonstrating the advantages of the knowledge-based interface and for maintaining and expanding the system.

Upon completion of the proposed additional research, implementation in a manufacturing environment would be possible. Certainly, this would be the ultimate test and would be necessary for general acceptance.

REFERENCES

1. General Motors Manufacturing Automation Protocol: A Communication Network for Open Systems Interconnection, 1986, p.A6-9.

2. Honeywell IPC-620 Communications Interface Module: Operator Manual, Process Control Division, 1984.

3. NND Assembly Robot 88: Operations and Service Manual, ATP, Inc. - Advanced Technical Products.

4. International Standard 7498: Information Processing Systems-Open Systems Interconnection- Basic Reference Model, American National Standards Association, Inc., 1984.

5. Stallings, William,Data and Computer Communications, Macmillan Publishing Company, 1985.

Appendices can be obtained from

Universal Energy Systems, Inc.

FINAL REPORT

DROPLET SIZE DISTRIBUTION MEASUREMENTS IN A SINGLE ELEMENT LIQUID ROCKET INJECTOR

Principal Investigator:

William M. Grissom

Asst. Professor of Physics

December 1987

Submitted in completion of requirements on :

Contract No. F 49620-85-C-0013/535851-0360
Purchase order No. S-760-GMG-040
Administered by : Universal Energy Systems, Inc.
4401 Dayton-Xenia Rd.
Dayton, Oh 45432

Sponsored by : Air Force Office of Scientific Services
Bolling AFB, D.C.

ABSTRACT

A coaxial, dual beam single particle counter employing absolute intensity measurements at 90 degrees has been developed for droplet size measurements in a dense single element rocket injector spray. A small diameter beam is used to locate a droplet in the center of a main beam. Preliminary testing using two laser beams of different polarizations has resulted in poor separation of the scattering signals. The system is presently being modified to a two color system. A simple technique termed "scattering tomography" has been developed to account for the attenuation of the laser beam in passing through the spray. The fluctuations in the beam intensity caused by the spray are also accounted for.

INTRODUCTION

In modeling the spray combustion and combustion instability problems in a liquid rocket engine the parameter which has perhaps the greatest influence and is also the least well-known is the droplet size distribution. The results obtained by present numerical combustion models are a strong function of the assumed size distribution. In the real world environment, small changes in the injector design can significantly change the rocket performance. In order to develop a more systematic design methodology for liquid rocket engines, more exact measurements of the droplet size distributions are required.

A number of techniques have been developed to determine droplet sizes in sprays [1]. A number of these techniques have been applied with limited success to rocket injector sprays. The earliest technique employed hot wax sprays. Much of the injector design data used today is based on these results. The difficulties in properly scaling the results for the different fluid properties of the wax make this data unreliable for determining the absolute droplet sizes. The data is mainly useful for understanding the dependance of droplet size on various injector parameters.

Later research attempted to measure the droplets photographically, either directly or through intermediate holograms. These suffer from difficulties in determining satisfactory focus criteria and in processing the volumous data needed to obtain a statistically accurate sample size. Other problems are that there is bias towards slower-moving droplets, since the technique gives a spatial average. For these reasons photography is rarely attempted today. Most techniques presently considered employ laser scattering in one of several forms which are itemized below.

Diffraction pattern, time-averaged:

This technique is presently implemented in an instrument marketed by Malvern Inst.,Ltd. This method is most sensitive for smaller droplets. It gives a droplet size distribution weighted over a line-of sight through a spray. It also gives a spatially-averaged size distribution rather than the desired temporal average. Because the signal is a DC level there is no way to discriminate the signal from the background glare caused by multiply scattered light. Indeed, this has been found to cause a significant bias for sprays with large beam attenuations [2,3]. While this techniques is useful for studying gross changes in spray charateristics it is doubtful that it can give sufficient details to adequately characterize injector designs.

Laser-velocimeter based systems :

The earliest of these attempted to correlate the visibility measured in forward scattering to droplet size [4]. However, the visibility technique has never proven satisfactory due to the difficulty in unambiguously relating the visibility to droplet size.

One presently investigated system utilizes the phase shift of the doppler signal at different collection angles [5]. While this system has performed well in some sprays its application to the very dense rocket injector sprays is questionable. The doppler signal is much weaker than the overall intensity signal and is of a higher frequency, making it difficult to distinguish from noise. In the phase shift system only a fraction of the total doppler signal is used. To maximize the signal forward scattering is normally used, which greatly increases the probe volume size. One might also expect the phase fronts of the laser beams to be disturbed more easily than the beam intensity by droplets in the spray, which would destroy the interference pattern at the probe volume.

Intensity Measurements:

The earliest of these simply measured the light pulse scattered when a droplet passed through a laser beam. The pulse height could be related to the droplet size [6]. The size of the probe volume is defined by imaging the laser beam onto a slit or pinhole. One problem with this techniques is that the laser beam intensity varies across the profile so that the pulse height depends upon the trajectory of the droplet through the beam. A similar problem is that a droplet can be partially obsured by the slit or pinhole. It is possible analytically to compensate the data for both of these effects [7], however it is preferable to eliminate them at the data collection stage. The present coaxial beam sizing instrument accomplishes this.

When the present design was first conceived the author was unaware of two similar designs which address this same problem. The first, designed by C.F. Hess [8], employs a laser velocimeter probe volume to locate a uniform intensity region within a larger beam of a different polarization or color. While this method has proved very satisfactory in sparse sprays, the same concerns about the detection of the doppler signal in dense sprays apply. The second design by J.C.F. Wang et al [9] is even similar to that suggested here, with the main exception that they measure the diffracted light scattered directly forward in a system designed for very sparse particle densities.

1. Optical Design

In the coaxial beam droplet sizer the droplet size is determined by the peak scattered intensity when a droplet passes through the center of the main laser beam. A smaller diameter locator beam is concentric with the main beam and serves to define the center of the main beam. The locator beam is of a different polarization or color in order to discriminate the scattered signals from each beam.

The scattered light is collected at 90 degrees and imaged onto the main slit 300 microns wide, which defines the probe volume size. This light is then collimated and the light from the main beam is reflected away to a photomultiplier by a polarizing beam splitter cube. The remaining light, due to the locator beam, is refocused at a magnification of x10 onto the locator slit which passes the center 150 microns of the main slit to a photomultiplier. This ensures that droplets up to 100 microns large are not obscured by the edges of the main slit. This is depicted in Figure 1.

Probe Volume Size:

As a design basis a monodisperse spray of 80 micron diameter droplets with a mean spacing of 5 droplet diameters was assumed.

The locator beam diameter must be quite small at the probe volume to satisfastorilly define the center of the main beam. However, if it were so small along the entire path through the spray a single droplet could completely scatter it. To avoid this the locator beam is first expanded to a large diameter and then steeply converged to the probe volume.

One would expect to first encounter a spray droplet of diameter d at a distance of one mean free path from the probe volume:

$$l = (5d)^{**}3 \; / \; [\pi/4 \; D^2]$$

where: D = beam diameter at l

In terms of the f no. of convergence:

$$f = 159 \; (d/D)^{**}3$$

The fraction of the beam scattered by this closest droplet is:

$$e = (d/D)^{**}2 = (f/159)^{**}2/3$$

In the present design the locator beam in converged at f/6.5 giving e = 12 %, which should result in little effect upon the locator beam by the closest droplet, on average.

From Gaussian optics the locator beam diameter at the probe volume can be calculated:

$$Do = 4\lambda f/\pi = 5.2 \text{ microns}$$

The useable length of the locator beam at the probe volume is 0.65 mm which is the distance over which it is less than 50 microns in diameter. This is fairly easily aligned.

A 20 x microscope objective was used to expand the locator beam to 40 mm diameter. A spatial filter assembly was used, however the pinhole was found to be unneccessary. A pair of achromatic lenses were used for both the focusing and collection lenses. The main beam is brought coaxial with the locator beam by a small 6 mm diameter mirror inserted in the center of the locator beam

In choosing the main beam diameter several conflicting aspects were considered. The advantage to making the main beam larger is that the intensity is made more uniform in the center region and the beam is less affected by droplets along the path through the spray. Where as the disadvantages are that the lilklihood of having more than one droplet in the probe volume at a time increases, and the beam intensity decreases. A droplet could be dispaced up to 85 microns from the center of the probe volume and still trigger the locator beam. To maintain the intensiy ± 10 % the main beam diameter must be at least 370 microns.

While it is difficult to provide for a wide sizing range in a dense spray with any system, the present design is easily adjusted to optimize the performance by changing the main beam diameter and the slit widths.

2. Signal Processing

Two high gain (2.E7) photomultipliers were initially selected. However the main beam current pulses were so large they exceeded the linear range of the standard dynode resistor chain used. To maintain linearity a lower gain (6.E5) tube was used for the main beam with a specially wired high current (4ma) dynode chain employing zener diodes and capacitors on the last 3 stages. There was no concern with linearity for the locator detection.

A droplet could have a velocity as high as 100 m/s and could cross the 300 micron main beam in 3 microseconds. This required electronics with frequency response flat to 170 kHz.

To obtain a voltage pulse the photomultiplier output was terminated across a resistor. Assuming 100 pF capacitance in the tube and cable, a resistor as large as 2000 ohm could be used without attenuating the signal. The negative voltage pulse so obtained was amplified by a homemade inverting op amp circuit (300 kHz). A LeCroy 612A amp, with the termination increased to 1000 ohm, was used as a preamp when needed. The resulting positive voltage pulse was found to have significant high-frequency noise, which would confuse the pulse- height analyzer. To smooth the pulse it was passed through a homemade active 4th-order low-pass filter set at 150 kHz.

A single channel pulse-height analyzer (Ortec 550) was used. The output pulses were sent to an "and" gate, with the locator beam signal as the other input, to insure that only droplets passing through the center of the main beam were counted. The output pulses from the "and" gate were averaged by a ratemeter and sent to an X-Y plotter. The plotter was slowly

swept across the pulse-height spectrum. The sweep signal was amplified and used to sweep the lower level of the pulse-height analyzer. This is shown in Figure 2. In cases were the data rate was very low a digital counter was used instead.

3. Data Analysis

For droplets larger than 10 microns the Mie oscillations due to interference should average out over the collection angle resulting in the pulse heights being proportional to the square of the droplet diameter, for constant refractive index. A single calibration point would then suffice. However a number of corrections to the data are neccessary as itemized below.

a. Beam attenuation by the spray

This has been mentioned as the chief drawback to using the absolute intensity to determine droplet size, however it is fairly simple to determine the beam intensity at the measurement point by passing the beam to the point alternately from the left and right side of the spray as shown in Figure 3.

Assume that the beam attenuation along any path is the same for either beam direction. Then,

$$I1/Io = (It/Ir)$$

expanding:

$$I1/Io = (It/Io)(I1/Ir)(Io/I1)$$

Finally,

$$I1/Io = \mathrm{sqrt}[(I1/Ir)(It/Io)]$$

Note that no explicit assumption as to the absorption law or symmetry of the spray is required.

The pulse-height distribution measured with the beam from each direction should be the same shape, with the abcissas simply scaled according to the intensity ratio. Any quantity in the distribution proportional to the abscissa, such as the mean height or area, can be ratioed to determine I1/Ir. The ratio It/Io can be taken simply as the ratio I1/Ir at the edge of the spray (or anywhere the incident intensity can be directly measured) or It can be measured directly using a photosensor, using care to subtract the multiply scattered light.

Attenuation of the collected light in passing out of the spray must also be considered. It can be argued that the collected light will be attenuated by the same fraction as the laser beam along the same path, allowing the data to be compensated accordingly. This is discussed further in Sec. 3d.

In physically performing these measurements it may be easier to reorient the spray than to actually pass the laser beam from each side. To maintain the exact same geometry the spray would need to be rotated 180 degrees about the horizontal collection axis. However for sprays with assumed symmetry about two orthogonal axes, such as a like-doublet injector, the spray could be rotated about the vertical axis through the collection point, if the laser beam passed through the center of the sprsy

as shown in Figure 3. For each of these geometries the collected light would be attenuated by the same fraction so that the ratio Il/Ir would be unaffected.

This determination of the beam intensity from scattering measurements will be termed "scattering tomography" to denote the similarity to the much more involved absorption tomography technique. The present technique can be simply implemented without requiring any additional spatial measurements than would normally be taken in a complete spray analysis.

b. Fluctuations in beam intensity

The previous section considered means of correcting the data for the average attenuation of the beam at the measurement point. A related concern is that there is a statistical fluctuation in the beam intensity due to variations in the number of attenuating droplets along the path.

In order to calculate the beam attenuation it is assumed that a droplet scatters away all of the light incident upon it. Diffraction is not considered since virtually all of the diffracted light will remain within the beam boundaries over the short paths in an injector spray. It is further assumed that the beam profile is smoothed out downstream of a droplet. Both diffraction and light refracted through the center of a droplet will tend to accomplish this. Shadows in the beam profile caused by upstream droplets are not considered here and would make the actual fluctuations greater than calculated .

The fraction of the beam scattered by a droplet is (d/D)**2 where d is the droplet diameter and D is the exp(-2) beam diameter. Using Beer's law the fraction of total beam intensity attenuated is: A = 1 - exp[-N(d/D)**2] ; where N = no. droplets. According to Poisson statistics this attenuated fraction will fluctuate by a fraction sqrt(N) / N rms, giving a fractional fluctuation of:

$$Irms = A(d/D)[1/\sqrt{[-\ln(1-A)]}]$$

For both very low and very high attenuations the fluctuation goes to zero. The maximum fluctuation occurs at an attenuation of 71.5% in which case Irms = 0.638 (d/D). The present design case of 80 micron droplets in a 300 micron beam could have a maximum intensity fluctuation of +/- 17% , corresponding to a diameter fluctuation of +8%, -9% for a monodisperse spray. This could be reduced by converging the main beam to the probe volume to decrease the sensitivity to individual droplet.

Hess [10] measured diameter fluctuations of +7%, -10% with 5 % attenuation using 73 micron droplets in a 1000 micron (?) main beam. The above analysis would predict only +/- 0.8%. The difference may be due to shading, however it is also possible that his LDV visibility trigger signal was affected by the spray. In either event , if the intensity fluctuations can be either calculated or measured then the effect upon the size distribution is known and compensation can be made.

A related concern is fluctuations in the pulse height due to statistical fluctuations in the number of photons collected. However, this does not appear to result in significant noise. The calculations are given in Appendix II.

c. Count rate statistics

With most single particle counting systems it is neccessary to compensate the measured count rates for the probability of a certain size droplet entering the probe volume and triggering a measurement. In the case of a very large probe volume all droplet sizes would have equal probability of being found in the probe volume. However, in a dense spray a minimal probe volume is required, making this correction significant.

In the present case a measurement is triggered when a droplet crosses the locator beam. One cannot actually speak of a droplet being outside of the locator beam since the intensity falls off exponentially from the center. Assume that the trigger threshold is set so that any droplet not crossing the exp(-2) radius does not trigger a measurement and that any droplet crossing this radius does trigger the measurement, regardless of size. This is a reasonable assumption since the locator beam is so small (5 microns) that the full area of a droplet does not significantly affect the total energy scattered. Experimental measurements show the locator pulses to be very large so we might readily assume that whenever a part of the droplet is in the view of the locator slit it will trigger a measurement.

With the above assumptions a measurement would be triggered whenever the center of a droplet is roughly within a cylinder of diameter D + d and length S + d ; where: D1 = locator beam diameter, d = droplet diameter, and S = slit width. The relative probabilities of a droplet of diameter d1 triggering a measurement to one of diameter d2 is given by the ratio of their trigger volumes:

$$\frac{p1}{p2} = \left(\frac{S + d1}{S + d2}\right)\left(\frac{D + d1}{D + d2}\right)^2$$

There are a number of simplifications in this derivation. In actual fact the probabilities should be biased more towards larger droplets. The relative probabilities can be determined experimentally by mixing known volumes of two sizes of uniform glass spheres. The difficulty then is in insuring equal flow velocities and distributions in a flowing system. The author is considering a simple arrangement whereby the spheres are placed in an evacuated flask and allowed to fall through the probe volume. In this "batch" technique the results would not be biased by velocity. It is only neccessary to assume that the droplets are spread randomly.

d. Multiply scattered light

A final concern with the laser beam attenuation is that the attenuated light is not lost but rather is multiply scattered between the droplets. No general solution to the multiple scattering problem has been found.

The largest contribution of multiply scattered light is a steady background glare against which the droplet is seen. As mentioned, this causes a significant bias in cases where a DC signal is measured as with the Malvern diffraction system. However, with single particle counting systems a pulse signal is measured. The multiply scattered light primarily causes an increase in the baseline level which is fairly easily eliminated.

In determining the effects of multiple scattering upon the present technique, the proper question is whether the presence of the droplet in the probe volume would change the amount of multiply scattered light directed into the collection angle. At high attenuations the multiply scattered light is almost isotropic so that a droplet could not send it in a preferred direction. At lower attenuations the multiply scattered light is travelling more in the forward direction so that it could give a net contribution to the laser beam, however simple energy considerations would argue that for attenuations less than 90% it should not cause a significant contribution to the laser beam. In any event the scattering tomographic procedure discussed previously would at least partially account for any multiple scattering contribution to the laser beam.

A final consideration is whether the assumption that the collected light would be attenuated by the same fraction as the laser beam over the same path is true. Initially one might imagine that the lens would collect much of the scattered light since it has such a large collection angle. However, for this scattered light to recontribute to the image it must be turned to a specific angle. This is identical to the requirements for scattered light to recontribute to the laser beam and should have the same probability.

4. Progress to Date

Due to delivery problems not all of the equipment was received until 8/1. System testing began about 8/20. As of this writing several problem areas have been investigated. These are discussed here.

The greatest problem to date is that the polarizations of the two beams cannot be adequately separated. Although initially conceived as a two-color system the polarization approach was adopted for cost reasons, following the approach of Hess [11]. Hess has had similar difficulty in separating the two polarizations.

With a little thought it is clear how a droplet can change the polarization of a ray. A vertically polarized ray which is refracted upward would have the polarization rotated towards the horizontal. Indeed this phenomenon has been used as a diagnostic

technique by Beretta et al [12], as the polarization ratio depends upon the droplet size for droplets smaller than 20 microns.

Fortuneately, in the last year low cost He-Ne lasers of various colors have become available. Presently the system is being converted to a two color system for which the beam separation is easily done with interference filters. Indeed, the depolarizing effect may prove useful in decreasing the sensitivity to the refractive index by proportioning the two polarization components.

For system checkout glass microspheres have been used. These are circulated in a circular flask by an air jet. The volume of air injected must be small to minimize carryover of spheres past the cotton filter. While convenient, this method has drawbacks. It is difficult to move the spheres fast enough so that the pulse-height analyzer works optimally. A related problem is that the count rate is quite low. The problem is that the fast moving air jet is devoid of spheres. The best results are obtained using a small pulsating air pump which tends to establish a tangential resonance in the circular flask.

A final problem is that the current pulses are not always smooth, Gaussian shapes. This occurs even at low densities so it is not due to beam attenuation. It is likely due to the photomultiplier shielding or dynode chain, which has been redesigned.

5. Future Plans

Due to the problems itemized it has not been possible to fully characterize the system performance, however most of these problems are being solved as of this writing. While the funds have been expended the remaining performance studies can be continued at marginal cost.

In order to allow glass microspheres to be used for convenient calibrations, yet perform measurements of spray droplets with different refractive index, it is neccessary to analytically determine the dependance of scattering intensity on the refractive index. A ray tracing code will be developed for this.

For marginal cost selected measurements in a spray can be made as a verification. However, to perform a full spatial characterization of a spray would require additional funds for a computerized translation system and data recording. Indeed, the ideal arrangement would replace the single channnel analyzer with a digital O-scope system which would provide multi-channel analysis of the pulses and allow determination of the droplet speed from the pulse width.

REFERENCES

1. B.J. Azzopardi; Int J Heat and M Trans, Vol 22, p1245 (1979)
2. A.J. Yule et al; 18th Symp (Int) on Comb, p1501 (1981)
3. L.G. Dodge; Opt Eng, Vol 23 no 5, (Sept/Oct 1984)
4. W.M. Farmer; Appl Opt, 13, p 610 (1974)
5. W.D. Bachalo, et al; AIAA-86-0296 (Jan 1986)
6. N.A. Chigier, et al; 17th Symp (Int) on Comb, p315 (1979)
7. Y. Muzutani, et al; Comb and Flame, 44, p85 (1982)
8. C.F. Hess; AIAA-85-1443, (July 1985)
9. J.C.F. Wang et al; Appl Opt, 25, no 5, p653 (March 1986)
10. C.F. Hess; Appl Opt, 23, no 23, p4375 (Dec 1984)
11. F. Beretta et al; Comb Sci and Tecch, 36, p19 (1984)
12. W.M. Grissom; A Feasibility Study of Liquid Rocket Engine Combustion Diagnostics, AFRPL-TR (1986)

Appendix I. Rocket Engine Injector Spray Characteristics

In a large, multi-element injector spray there is little radial spreading so that the droplet density is readily calculated from conservation of mass. Assume a monodisperse droplet spray with spacing 1 between droplets in a square lattice. The volume flow rate is:

$$\text{Vol.} = \frac{\pi}{6} \text{Vd} \left(\frac{d}{1}\right)^3 A$$

where: Vd = vel. of droplets
A = area of injector face

This equals the volumetric flow through the orifices in the injector:

Vo Ao ; where: Vo = axial vel. of fluid jet
Ao = total orifice area along face

Then the spacing between droplets is:

$$1/d = [\pi/6 \ (Vd/Vo)(A/Ao)]^{1/3}$$

We might take Vd = Vo. In fact calculations show that the droplets decelerate to about 0.8 Vo as they enter the combustion chamber and then are accelerated by the expanding combustion gases. For one representative injector ("Star-Tech") A/Ao =121, giving: 1 = 4.0 d. Other injectors are less densely packed, so that 1 = 5 d is a reasonable design basis.

The fraction of a laser beam transmitted through an injector spray is found using the formulation of section 3b. The droplet density is one per 1**3. The number of droplets along a pathlength L which a beam of area Ab contacts is then

N = L Ab / 1**3. Each droplet attenuates a fraction
e = π d**2/(4 Ab) of the beam, so that the total fraction transmitted is:

$$Ft = \exp(-e\,N) = \exp(-a\,L)$$

where: a = absorption const. = 1.5/d (Vo/Vd)(Ao/A)

Note that the mass flow rate of the spray is unimportant.

A uniform drop size has been assumed here. Since each droplet absorbs a certain fraction of the beam, regardless of its position in the spray, for polydisperse sprays the diameter above can be replaced with a reciprocally-weighted average diameter:

$$1/davg = (\text{Sum } 1/d)/N$$

This average is biased towards the smaller droplets which are generally considered less important in the combustion process. Hence beam attenuation is not particularly useful as a drop sizing technique.

One would expect a reciprocal averaged size of about 20 microns in an injector spray. This would give an absorption constant of 0.775/mm for the Startech injector. There is a 3.22 mm spacing between elements, so that even in a single element at the distance where the elements begin to mingle one would expect only 8.3% transmission. For this reason measurements close to the spray are possible only in a single element injector spray.

Some experimental measurements were performed by the author and Michael Powell at AFAL [12] on a 3.5 inch injector. Only in the case were the flow rate was inadvertently set 11 times too low was the output intensity measureable, in that case the reciprocal-weighted diameter was calculated to be 173 microns. At the normal flow rates no output beam could be measured. In this case the beam was seen to "bloom" into a spherical shape about halfway across the spray and no output beam was detectable. This is interpreted as the beam changing from a small, directed beam into an isotropic glare due to the multiple scattering.

Appendix II. Photoelectron Statistics

The current pulse obtained when a droplet crosses the main beam is actually composed of thousands of discrete pulses, each due to the release of a photoelectron from the cathode of the photomultiplier. Due to the system time responses and the purposeful low-pass filtering imposed, these discrete pulses partially merge into the final smoothed pulse with a ripple caused by the statistical fluctuation in numbers of photoelectrons, termed "photon shot noise", which is calculated below.

For the 5 mW, 300 micron laser beam the light scattered by an 80 micron droplet would be 0.31 mW. The f/5.2 lens collects 0.23% of the scattered light. Assuming that the scattering at 90 degrees is only 10% of the isotropic value, the power collected is 7.1 E-8 W. With 30% transmission through the optics and a photomultiplier quantum efficiency of 5%, the photo-electron rate is 3.4 E9 ph/sec. This corresponds with experimental measurements.

Over a 3 microsecond time frame 10,200 photoelectrons would result. According to Poisson statistics this would result in a variation sqrt N or +/- 1% rms. Thus a 5 mW laser is the minimum tolerable. Indeed, a larger larger laser would be neccessary if there were appreciable beam attenuation. For slower droplets more filtering could be used and these fluctuations would decrease.

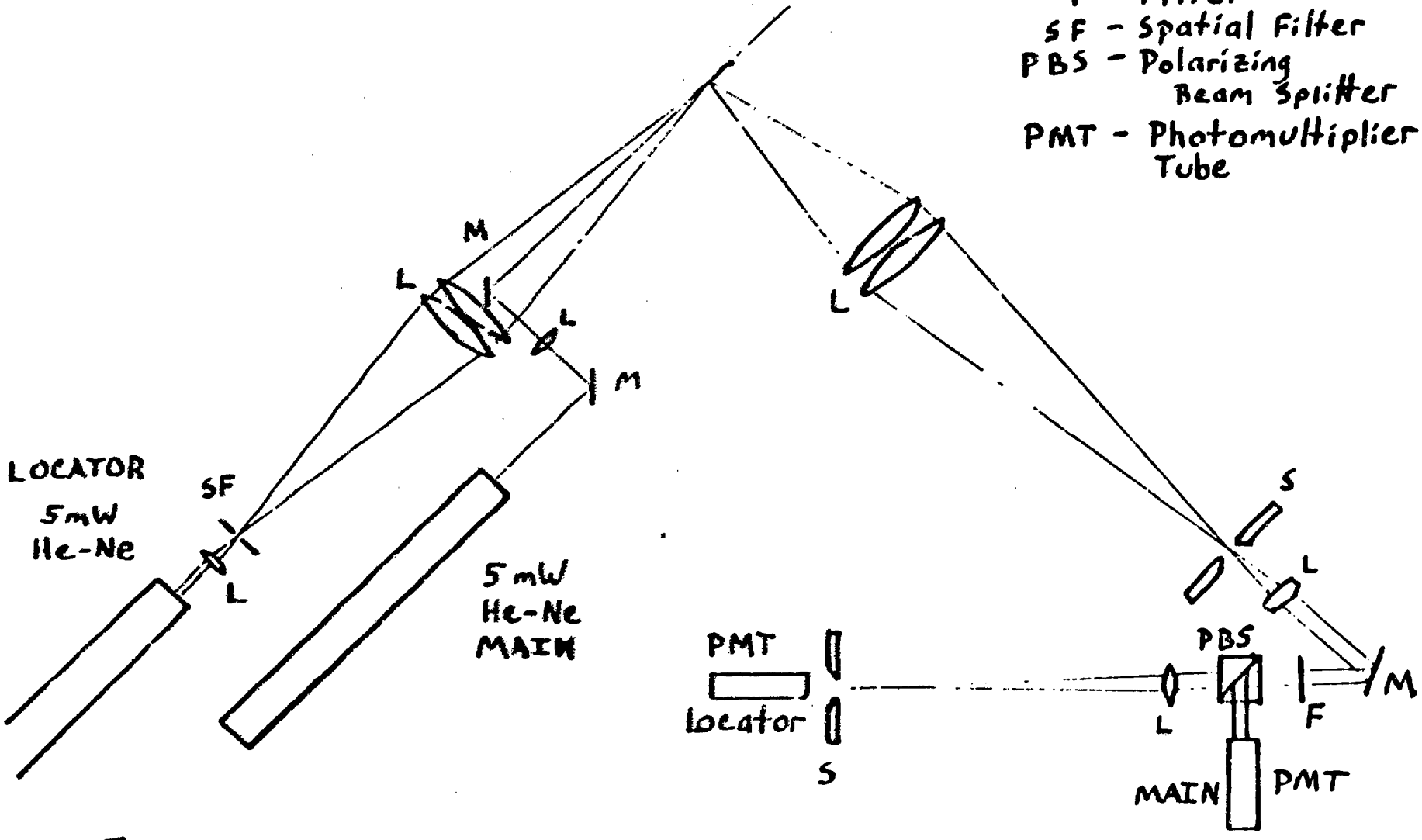


Figure 1. Optics

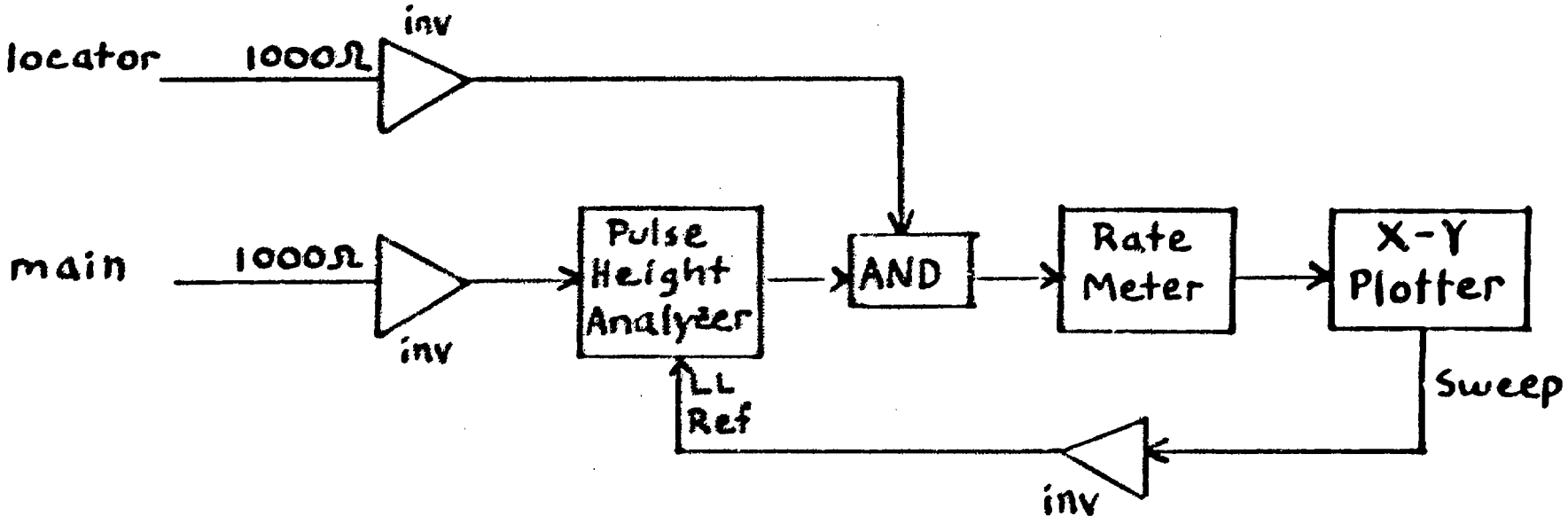


Figure 2. Signal Processing

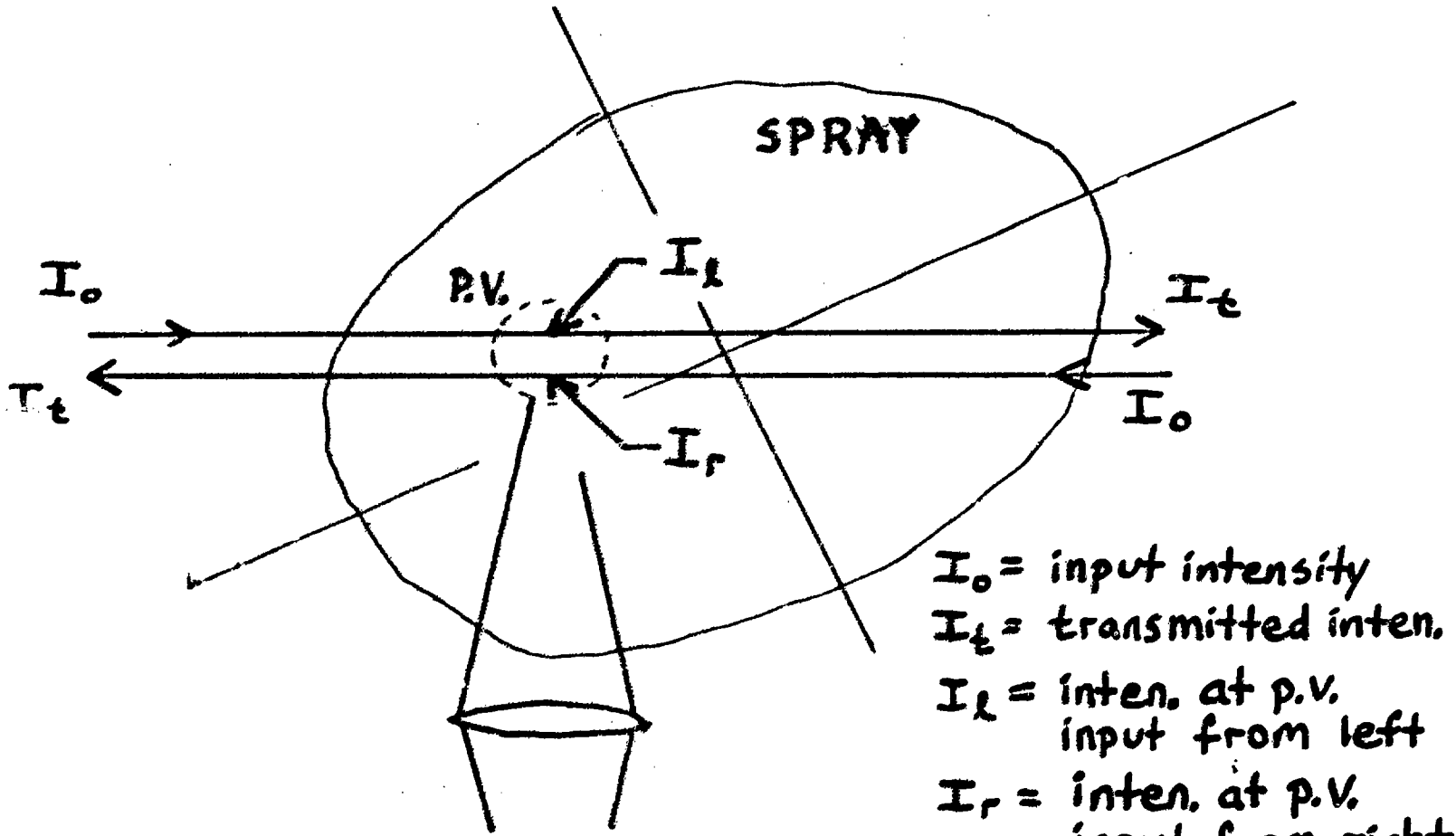


Figure 3. Scattering tomographic geometry